BIOPHYSICAL CHARACTERIZATION OF THE SITE – SPECIFIC BINDING OF INTEGRASE TN916 TO ITS COGNATE DNA

DISSERTATION

zur Erlangung der naturwissenschaftlichen Doktorwürde (Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich

von

Stoyan G. Milev

aus Bulgarien

Promotionskomitee

Prof. Dr. Hans Rudolf Bosshard

PD Dr. Ilian Jelesarov

Zurich 2004

Parts of this dissertation were published in:

1. Milev, S., A.A. Gorfe, A. Karshikoff, R.T. Clubb, H.R. Bosshard, and I. Jelesarov, *Energetics of sequence-specific protein-DNA association: conformational stability of the DNA binding domain of integrase Tn916 and its cognate DNA duplex.* Biochemistry, 2003. **42**(12): p. 3492-502.

2. Milev, S., A.A. Gorfe, A. Karshikoff, R.T. Clubb, H.R. Bosshard, and I. Jelesarov, *Energetics of sequence-specific protein-DNA association: binding of integrase Tn916 to its target DNA.* Biochemistry, 2003. **42**(12): p. 3481-91.

3. Milev, S.; Bosshard, H. R.; Jelesarov, I.; *Enthalpic and Entropic Effects of Salt and Polyol Osmolytes on Site-Specific Protein-DNA Association: The Integrase Tn916-DNA Complex.* Biochemistry, 2005. 44(1): p. 285-293.

4. Milev, S., Bosshard, H. R. and I. Jelesarov (2004). *Energetics of Site-Specific DNA Recognition by Integrase Tn916.* In Biocalorimetry: Applications of Calorimetry in the Biological Sciences. 2nd Edition. (Ladbury, J. E. and Doyle, M., eds). John Wiley & Sons, in press

SUMMARY

Protein-DNA recognition is a crucial link in the chain of molecular events leading to the realization of the information stored in the genes. This study presents a comprehensive thermodynamic and kinetic analysis of site-specific protein-DNA recognition. The system studied is the complex formed between the bacterial integrase Tn916 and its DNA target, a 13 base-pair DNA duplex. The N-terminal DNA-binding domain of the integrase (Int-DBD) positions the face of a three-stranded antiparallel β-sheet in the major groove of the DNA. This is a rare structural mode of DNA recognition.

Binding-induced conformational changes are often observed in macromolecular association reactions. Moreover, the thermal motions and the so-called soft vibrational modi of the interacting macromolecules can be changed considerably in the complex. The energetic costs of structural adaptation are significant but difficult to quantify. The first and major part of the present work (Chapters 4-6) deals with the contribution of structural rearrangements, thermal fluctuations and hydration effects to the net thermodynamic profile of the interaction of Int-DBD with its target DNA. To this end, a comprehensive analysis of the single components of the complex was indispensable. From a combination of temperature and denaturant induced unfolding experiments we find that Int-DBD is compact and unfolds cooperatively, showing only a small deviation from two-state behavior (folded/unfolded). The DNA target displays the typical thermodynamic signature of short duplexes: gradual, non-cooperative structural changes precede the sharp cooperative unfolding leading to strand dissociation.

Protein-DNA binding was studied by isothermal titration calorimetry (ITC), differential scanning calorimetry (DSC), circular dichroism (CD) and fluorescence spectroscopy. The heat capacity change accompanying the association reaction (ΔC_p) is temperature-dependent, decreasing from -1.4 kJ K^{-1} mol^{-1} at 4 °C to -2.9 kJ K^{-1} mol^{-1} at 30 °C. The reason is that the partial molar heat capacities of the free protein, the free DNA duplex and the protein-DNA complex do not change in parallel with the temperature increase and that thermal motions of the protein and the DNA are restricted in the complex. Measured values of ΔC_P are larger than the semi-empirical estimate of ΔC_P of -1.2 kJ K^{-1} mol^{-1}. The discrepancy between measured ΔC_P and structure-based estimates is due to conformational changes of the complex constituents accompanying the binding reaction and to incomplete dehydration of polar groups in the complex. In support, we observe cavities at the complex interface that are large enough to accommodate about 10 water molecules.

A macromolecular association process is sensitive to changes of the chemical composition of the environment. The effect of small compounds on the binding equilibrium can be used to explore the energetics and molecular mechanism of macromolecular interactions. In Chapter 7 we investigate the influence of NaCl and glycerol (an osmolyte) on sequence-specific protein-DNA association. Increasing the Na^+ concentration from 0.12 M to 0.32 M weakens the binding affinity by twenty-fold. Calorimetric experiments show that the net increase in binding free energy is caused by an increase of enthalpy and a decrease of entropy. We discuss this finding in the light of classical and modern views about the general salt effect. Addition of up to 30% glycerol leaves the binding affinity unchanged. However, the energetic partitioning of the free binding energy into binding enthalpy and binding entropy differs in aqueous and glycerol-containing buffer. We propose that the observed favorable enthalpic contribution and the unfavorable change in entropy are a consequence of the general hydration properties of polar and nonpolar groups, in combination with tighter packing in the presence of glycerol. Addition of salt lowers significantly the rate of protein-DNA association while the rate of dissociation is almost unaffected. In the presence of glycerol both kinetic constants decrease in parallel, leaving the net binding free energy unchanged.

To explore the role of individual protein-DNA contacts for the energetics and dynamics of the Int-DBD-DNA complex, several residues in hydrogen-bonding distance to DNA bases or backbone phosphates were substituted by alanine. The results presented in Chapter 8 clearly indicate that removal of such side chains weakens binding. Partitioning of $\Delta\Delta G$ (defined as $\Delta G_{mutant} - \Delta G_{wildtype}$) into enthalpic and entropic contributions reveals a complicated picture. There is no correlation between $\Delta\Delta G$ and the mutation-induced changes of ΔH and $T\Delta S$. Generally, the mutations change ΔC_P of DNA binding. Some mutations change the stability and the conformation of Int-DBD. This was indicated by pK_a shifts of some ionizable groups in the complex as compared to the free protein. No such shifts of pK_a values were observed when the wildtype protein bound to the DNA. Kinetic experiments showed that, generally, electrostatic interactions increase the rate of association: replacement of charged side chains by alanine slows binding whereas replacement of uncharged groups has no effect. In the case of the mutant Lys21Ala, replacement of charged lysine by uncharged alanine decreases the rate of association and increases the rate of dissociation. Stabilization of protein-DNA interaction by hydrogen bonds between the side chain of tyrosine 40 and A20 and C21 bases results from a slower rate of dissociation of the complex.

ZUSAMMENFASSUNG

Protein-DNA Erkennung ist ein entscheidendes Glied in der Kette der molekularen Vorgänge, die zur Verwirklichung der in den Genen gespeicherten Information führt. Diese Arbeit präsentiert eine umfassende thermodynamische und kinetische Analyse einer sequenzspezifischen Protein-DNA Erkennung. Es ist die Bindung der bakteriellen Integrase Tn916 an ihr Zielmolekül, eine 19 Basenpaare lange, doppelsträngige DNA. Die N-terminale DNA-bindende Domäne der Integrase (Int-DBD) bindet mit der Fläche eines dreisträngigen, antiparallelen β-Faltblatts in die weite Furche der DNA Doppelhelix, eine seltene Art der Protein-DNA Erkennung.

Bei makromolekularen Bindungsreaktionen beobachtet man oft bindungsinduzierte Konformationsänderungen. Zudem können die thermischen Bewegungen und die so genannten "weichen" Vibrationsmodi der interagierenden Moleküle in einem Komplex beträchtliche Veränderungen erfahren. Die energetischen Kosten der strukturellen Anpassung sind beträchtlich aber schwierig zu quantifizieren. Der erste und wichtigste Teil der vorliegenden Arbeit (Kapitel 4-6) befasst sich mit der thermodynamischen Bedeutung von struktureller Umordnung, thermischen Schwankungen und Hydratationseffekten für das thermodynamische Profil des Int-DBD-DNA Komplexes. Dazu war eine umfassende Analyse der einzelnen Komponenten dieses Komplexes unverzichtbar. Thermische und chemische Denaturierungsexperimente zeigten, dass das Protein Int-DBD kompakt ist, ein kooperatives Faltungs- und Denaturierungsverhalten aufweist und sich im Wesentlichen als Zwei-Zustandssystem (gefaltet/ungefaltet) verhält. Der DNA-Doppelstrang mit der Erkennungssequenz verhält sich wie andere kurze DNA Doppelstränge: man beobachtet allmähliche, nicht-kooperative strukturelle Veränderungen vor der abrupten Denaturierung, bei der die Stränge dissoziieren.

Die Bindung des Proteins an die DNA wurde mit isothermer Titrationskalorimetrie (ITC), *differential scanning* Kalorimetrie (DSC), Circulardichroismus (CD) und Fluoreszenzspektroskopie untersucht. Die Änderung der Wärmekapazität während der Bindungsreaktion (ΔC_p) ist temperaturabhängig; sie nimmt von -1.4 kJ K^{-1} mol^{-1} bei 4 °C auf -2.9 kJ K^{-1} mol^{-1} bei 30 °C ab. Der Grund liegt darin, dass die Änderung der partiellen molaren Wärmekapazität des freien Proteins, des freien DNA-Doppelstrangs und des Komplexes mit der Temperatur nicht parallel verlaufen und dass die thermischen Bewegungen des Proteins und der DNA im Komplex eingeschränkt sind. Die gemessenen Werte von ΔC_P sind grösser als der aus Strukturdaten berechnete semi-empirische Wert von -1.2 kJ K^{-1} mol^{-1}. Die Diskrepanz

zwischen den gemessenen und dem aufgrund der Struktur vorhersagten ΔC_P wird mit konformationellen Veränderungen der Komponenten im Komplex und mit der unvollständigen Desolvatisierung der polaren Gruppen im Komplex erklärt. Hohlräume in der Interaktionsfläche, welche etwa 10 Wassermoleküle aufnehmen können, bestätigen diese Interpretation.

Ein makromolekularer Bindungsvorgang reagiert empfindlich auf Veränderung in der chemischen Zusammensetzung der Umgebung. Der Einfluss von niedermolekularen Stoffen auf das Bindungsgleichgewicht gibt Aufschluss über die energetischen Eigenschaften und den Mechanismus makromolekularer Interaktionen. Im Kapitel 7 wird der Einfluss von Salz und osmotisch wirksamem Glyzerin auf die sequenzspezifische Protein-DNA Bindung untersucht. Der Anstieg der NaCl-Konzentration von 0.12 M auf 0.32 M erniedrigt die Bindungsaffinität um das Zwanzigfache. Kalorimetrische Untersuchungen zeigen, dass die Zunahme der freien Bindungsenergie auf der Zunahme der Enthalpie und der Abnahme der Entropie beruht. Diese Beobachtung wird im Lichte klassischer und moderner Theorien des allgemeinen Salzeffekts diskutiert. Die Zugabe von Glyzerin hat keinen Effekt auf das Bindungsgleichgewicht. Allerdings sind die Beiträge von Enthalpie und Entropie zur freien Bindungsenergie in wässriger und glyzerinhaltiger Lösung verschieden. Wir nehmen an, dass der günstige Enthalpiebeitrag und die ungünstige Entropieänderung eine Folge der generellen Wasserbindungseigenschaften polarer und apolarer Gruppen, in Verbindung mit einer festeren Packung in Gegenwart von Glyzerin, sind. Die Zugabe von Salz erniedrigt die Assoziationsgeschwindigkeit von Protein und DNA erheblich, während die Dissoziationsgeschwindigkeit gleich bleibt. Im Unterschied hierzu nehmen bei der Zugabe von Glyzerin beide Geschwindigkeiten gleichermassen ab, weshalb die freie Bindungsenergie gleich bleibt.

Um den Beitrag einzelner Protein-DNA Kontakte zur Gesamtenergie und zu den dynamischen Eigenschaften des Int-DBD Komplexes zu untersuchen, wurden mehrere Aminosäurereste, welche sich in H-Brückendistanz zu DNA-Basen oder zu Phosphaten des DNA-Rückgrats befinden, durch Alanin ersetzt. Die Ergebnisse, welche im Kapitel 8 vorgestellt werden, zeigen, dass der Ersatz der entsprechenden Seitenketten durch Alanin die DNA-Bindung schwächt. Die Aufschlüsselung von $\Delta\Delta G$, definiert als $\Delta G_{Mutante} - \Delta G_{Wildtyp}$, in enthalpische und entropische Beiträge gibt ein komplexes Bild. Es besteht keine Korrelation zwischen $\Delta\Delta G$ und der mutationsbedingten Änderung von ΔH and $T\Delta S$ zur freien Bindungsenergie. Tendenziell ist ΔC_P für die Mutanten kleiner als für den Wildtyp. Einige Mutationen verändern Stabilität und Konformation von Int-DBD. Ein klarer Hinweis hierfür ist die Veränderung des pK_a-Wertes von ionisierbaren Gruppen im Komplex verglichen zum freien Protein. Kinetische Experimente zeigen, dass elektrostatische Interaktionen im Allgemeinen die

Bindungsgeschwindigkeit vergrössern: Substitution geladener Seitenketten durch Alanin reduziert die Assoziationsrate, die Substitution ungeladener Seitenketten zeigt diesen Effekt nicht. Bei der Mutante K21A ist sowohl die Assoziationsgeschwindigkeit reduziert als auch die Dissoziationsgeschwindigkeit erhöht. Die Stabilisierung der Protein-DNA Interaktion durch Wasserstoffbrücken zwischen der Seitenkette von Tyrosin 40 und einer der beiden Basen A20 und C21 resultiert aus der langsameren Dissoziationsgeschwindigkeit des Komplexes.

1 INTRODUCTION

Studies of protein–DNA interactions are of fundamental importance for understanding the biological function of DNA. The information that is encoded in the genes is read and executed by proteins. Proteins are utilized as enzymes, recognition domains and other controlling and regulatory elements participating in chromatin packaging, replication, transcription, restriction, recombination and transposition. The prerequisite for all these processes is protein-DNA binding. Knowledge about the mechanism of the recognition provide unique insights into the way the living cell is functioning as a self-maintaining entity. Moreover, accumulated information on the structural, energetic and kinetic aspects of protein-DNA binding may eventually lead to the development of tools and strategies directed to practical and bio-medical applications and will most certainly facilitate studies of the organism on higher levels.

To date no universal structural or energetic code for protein-DNA recognition is known [8-13]. No firm correlations have been established between the chemical nature of amino acid side chains present at the DNA-recognizing protein face, the sequence of nucleotide base pairs, and the type of contacts they make. Only some loose and very general relationships have been found. The majority of the DNA-binding domains have a predominantly positively charged binding surface providing electrostatic complementarity to the negatively charged DNA molecule. However, it appears that in most of the cases studied so far electrostatics solely helps to speed up binding to the diffusion limit through long-range steering, while it contributes little, even opposes binding in terms of free energy [14-16]. The main driving force holding protein and DNA together is represented by the van der Waals contacts. Indeed, known structures of protein-DNA complexes exhibit a very good steric compatibility of the DNA-recognition protein motif with the geometric requirements of the major or minor DNA grove. Specificity and strength of protein–DNA complexes appear to be interrelated and to vary widely. In some cases binding is weak, unspecific and driven purely by electrostatic attraction. On the other extreme, binding can be very strong and completely intolerant to any change of the composition of the binding interface. High specificity is often achieved by conformational adaptation transitions of protein domains, or DNA duplex bending, or both [17-20]. Water molecules trapped at the protein-DNA interface frequently optimize the structural complementarity [21]. A research program combining structural, energetic and kinetic characterization of protein-DNA complexes will not only deepen our fundamental knowledge about the principles of macromolecular recognition but will also facilitate progress toward pharmaceutical and gene-

therapeutic applications aiming at modulation the function of protein-DNA complexes of interest, e.g. abolishing or strengthening binding, changing specificity, designing novel specificity, etc. In this context, understanding the nature and spatial determinants of the forces that drive and keep DNA-binding proteins and their cognate DNA together is of special biological, medical and industrial interest.

In this introduction some general structural features of protein-DNA complexes are first summarized briefly (section 1.1). Section 1.2 deals with the thermodynamic foundation of the protein-DNA recognition process. The emphasis is put on the informational content of observed free energy, enthalpy, entropy and heat capacity changes. Calorimetric methods as powerful tools for a detailed thermodynamic description of macromolecular binding reactions and macromolecular stability are introduced in section 1.3. Finally, general and experimental aspects of protein-DNA binding kinetics are discussed in section 1.4.

1.1 Structural features of the protein – DNA complexes

The structure of over a hundred protein-DNA complexes is known in atomic detail from x-ray crystallography and NMR spectroscopy. Nature utilizes various strategies for protein – DNA recognition. As in other macromolecular complexes, structural complementarity is a very important prerequisite for site-specific binding. Protein-DNA complexes bury a large amount of molecular surface. On average, the recognition interface spans an area of 1600±400 $Å^2$, with 12±3 base pairs and 24±6 amino acids involved in interactions [22]. A common feature is the higher polarity of protein-DNA interface as compared to the protein – protein complex interfaces. The majority of the protein-DNA complexes exhibit very tight shape complementarity. The interface packing density is estimated to be close or equal to that in the protein core [22, 23]. Obviously, tight packing of complementary shaped surfaces facilitates short-range interactions between the various groups and thus provides a major driving force for strong binding and slows down dissociation. It has been proposed that specificity of binding is achieved by two formally separate processes, namely shape recognition and chemical recognition [9]. For shape recognition, the architecture of the DNA-binding protein domain is important since the DNA molecule is structurally rather monotonous, while for chemical recognition, the determinants are chemical nature of the groups at the interface as well as their mutual orientation and charge compatibility.

The available structures of duplex DNA – binding proteins can be systemized in eight grand families according to the secondary structure motifs that are utilized for contacting the DNA [24]. Most commonly, the proteins insert an α-helix into the major groove of the B–DNA

duplex, but recognition modes involving β-sheets or loops are also occurring. One of the most widely spread motifs is the helix–turn–helix motif, which is present in numerous prokaryotic and eukaryotic transcription factors. Usually the motif comprises two almost perpendicular α-helices, one of which is inserted into the major groove of the DNA duplex. The largest grand family is the zinc finger motif in eukaryots. In these proteins zinc atoms coordinate His and/or Cys residues in α-helices and/or β-sheets to form finger shaped domains wrapping along the DNA molecule in the major groove. The zipper motif proteins are subdivided into two families that are present only in eukaryotic organisms: basic leucine zippers (b-ZIP) and helix-loop-helix zippers (b-HLH). The name is derived from the zipper-like way in which the two α-helices comprising the dimerizing domain interact with each other to form a superhelix. The duplex is recognized symmetrically in the major groove by two α-helices (one from each monomer), which follow the zipper domain directly (b-ZIP) or after a loop, the latter also contributing DNA contacts (b-HLH). The β-sheet family proteins usually utilize the edges of large (up to 10 stranded) β-sheets to contact the DNA in the minor groove. The β-hairpin/ribbon proteins differ from the β-sheet motif proteins by the generally smaller sized β-sheets (2 or 3 stranded). Depending on the family they can bind to the major groove inserting β-sheets lying flat inside or to the minor groove with amino acid side-chains intercalating between the DNA bases. Apart from the mentioned families, there is a large variety of structural strategies for recognition. Not only α-helices and β-sheets serve as structural binding unit. In some cases contacts from flexible inter-strand and inter-domain loops without regular structure achieve tight binding.

Shape recognition is thought to provide the “bulk” binding force, but the fine tuning for specificity is believed to stem mostly from chemical recognition. Certain common patterns of side chain-to-DNA contacts have been observed [9]. Arginine and glutamine residues having relatively long side-chains tend to interact with nucleotide bases, which are the only sequence specific part of the DNA molecule. Arginine is frequently found contacting the N7 and O6 atoms of guanine while glutamine contacts the N6 and N7 of adenine. The methyl group of thymine is often involved in van der Waals contacts with nonpolar amino acid groups. Most of the DNA binding domains are rather small and are thus unable to contact enough nucleotide pairs for achieving sufficient specificity. One way of overcoming this problem is employing two or more domains for recognition [25, 26]. Some major groove binders use loops or other flexible structural elements to contact the minor groove as well [27, 28].

It has long been recognized that water molecules play an important role in protein-DNA association [21]. Structural and computational analyses have identified water molecules bridging

protein and DNA groups, often participating in hydrogen bonding networks and serving as electrostatic screen between like charges [29]. In a few cases, like for example in the *trp* repressor-operator complex, no direct base-specific protein-DNA contacts are observed, and specificity is rendered through water-mediated polar contacts [30]. A survey of high-resolution protein complexes has identified on average one strongly bound water molecule per 150 $Å^2$ of protein-DNA interface [22].

One distinct feature of protein binding domains is their "adaptability". Many proteins are able to bind non-cognate sequences in a topologically similar manner as their cognate sequences by structural rearrangements. Such changes range from rotation of bonds to large-scale conformational transitions [31]. These perturbations in the interface are often costly in terms of energy and thus reduce the affinity. Frequently, structural elements or even entire domains fold simultaneously with DNA binding. One classical example is presented by b-ZIP proteins, in which the basic helices are partially or completely unfolded in the free protein but become fully structured in the major groove upon binding. The DNA duplex also plays a certain role for specific recognition. Since the DNA molecule is stiff, binding specificity is sometimes attributed to the flexible nature of the protein. However, DNA can adopt local conformations that differ depending on the sequence variation. For example, the hydrogen bond pattern in the major groove is different for AT and TA base pairs, as well as for GC and CG base pairs in the B-form duplex conformation but not in A-form [9]. Axial bending of the DNA is also important. Analysis of the structures of numerous protein–DNA complexes shows that proteins can bend the DNA duplex by different angles, sometimes severely, and such bending is crucial for formation of specific protein-DNA contacts. The propensity of the DNA to bend is also sequence dependent [32]. Both association-induced protein conformational transitions and DNA structural changes are linked to energetic expenditures that may significantly influence the energetics of binding. A recent analysis of the thermodynamic parameters of 10 protein–DNA complexes demonstrates an interesting correlation between the angle of duplex bending and the relative enthalpic and entropic contribution to the free energy change [18]. When bending is absent or small the favorable enthalpic contribution is large but the entropic penalty is large as well, as in the case of b-ZIP proteins. In cases where DNA structure distortion is pronounced, like for example the 87° DNA bending upon CAP protein binding, complex formation is entropy driven with substantial enthalpic penalty.

1.2 Thermodynamics of the protein – DNA interactions

From structural information alone it is not possible to draw quantitative conclusions about the stability of a protein-DNA complex and to predict how binding affinity will change in response to mutation or changes of environmental parameters, such as temperature, pressure, addition of low-molecular substances, etc. To find links between the sequence, structure and function of macromolecules is a long lasting endeavour in contemporary biochemistry and biophysics. Energetics and dynamics become more and more recognized as key pieces of information required for a comprehensive description of biological processes. Processes involving proteins and/or DNA are formally classified as folding transitions and binding reactions, but in general the underlying thermodynamic principles for these processes are fundamentally similar and often folding and binding are interrelated. Thermodynamics of the binding reactions describes a system of molecules in terms of temperature, pressure, free energy, enthalpy and entropy, thus giving us information of whether and why these reactions are going to occur. Thermodynamics has proved its power in understanding the behavior of simpler systems like gases, liquids and small molecule reactions, and develops more and more as a useful tool in the research of more complicated systems like proteins-DNA complexes.

Protein-DNA association is a noncovalent reaction comprising contributions from van der Waals contacts, oriented hydrogen bonds, salt bridges and long range electrostatic interactions, all of them with specific dependence on the environmental variables – temperature, pH, osmotic pressure and salt concentration. The binding often requires conformational adaptation of both the protein and the DNA molecule. Water and other solutes can be present at the complex interface, contributing to the overall energetic balance of binding as well. As a consequence, protein-DNA association is a dynamic process between flexible molecules and depends on a balance of many different forces. A thorough analysis of specific protein – DNA complex formation in terms of energetics would comprise measuring the stability of the single components and its changing with the temperature and the solvent composition; determining the basic energetic signatures of the binding reaction – the changes in free energy, enthalpy, entropy, heat capacity and how they vary with the temperature and the type of solvent; mutational analysis of the binding interface – the thermodynamic changes caused by removing residues important for binding; kinetics of the binding of the wild type and mutated protein to the DNA and its solvent dependency.

In the search for the unifying principles of the binding thermodynamics a challenging goal is to link the energetics of a protein-DNA association reaction with the type and number of

non-covalent intermolecular contacts at the binding interface, taking into account also the energetic consequences of structural adaptation and changes in hydration. The starting point is the determination of the changes of the free energy (ΔG), enthalpy (ΔH), entropy (ΔS) and heat capacity (ΔC_P) accompanying protein-DNA association.

1.2.1 Free energy changes

The most important thermodynamic parameter describing binding is the Gibbs free energy change (ΔG). It is the quantitative indicator of the equilibrium population of the different reacting species at given temperature, pressure and solvent conditions. For the simple equilibrium between protein, P, and DNA, D, and their complex, PD: $P + D \leftrightarrows PD$
the equilibrium constants is given by:

$$K_{eq} = \frac{[PD]}{[P][D]}$$

where the square brackets indicate the equilibrium concentrations. The free energy ΔG is related to K_{eq} through the fundamental thermodynamic equation:

$$\Delta G = -RT \ln K_{eq} \quad (1)$$

The free energy of the protein – DNA complex formation is the sum of enthalpic and entropic energy terms (eqn 2).

$$\Delta G = \Delta H - T\Delta S \quad (2)$$

Both components of ΔG depend on the temperature, so that eqn 2 can be written as:

$$\Delta G(T) = \Delta H(T_R) - T\Delta S(T_R) + \Delta C_P\left[T - T_R - T\ln\left(T/T_R\right)\right] \quad (3)$$

where ΔG(T) is the free energy change at any temperature T, T_R is an appropriate reference temperature, for which the enthalpy change, $\Delta H(T_R)$ and entropy change, $\Delta S(T_R)$ are known, and ΔC_p is the heat capacity change. Equation 3 predicts that ΔG changes with temperature. The function has an extremum at T_S, at which temperature ΔS = 0 ($\partial\Delta G/\partial T = -\Delta S$) and binding is driven by enthalpy only. Since the heat capacity change of protein-DNA binding is significant, the temperature dependence of the enthalpic and entropic terms are large and have opposite signs, leading to a much smaller temperature variation of ΔG. This phenomenon is known as temperature-dependent enthalpy-entropy compensation [33, 34] and is grounded in the fundamental thermodynamic laws for systems with non-zero heat capacity change. It can be shown that $\Delta H \approx T\Delta S + \Delta C_p(T_S - T_H)$, T_H being the temperature where ΔH = 0, and there is indeed a strong linear correlation between ΔH and ΔS since the second addend on the right-hand side of this expression is a constant for each system. The theory states that the temperature

dependent enthalpy-entropy compensation is strongest around T_S since the ΔG function goes through an extremum. Often T_S for biological macromolecular processes falls in the physiologically relevant temperature range and it appears, therefore, that enthalpy-entropy compensation in protein folding and binding is grounded in the specific chemical features of the macromolecular surface. However, the physical origins of this phenomenon are not fully understood. According to some views [35] the energetics of macromolecular processes is largely dominated by the energetics of water reorganization upon protein folding or protein-ligand binding. Formation of highly enthalpic water structures around hydrophobic surfaces at low temperatures is accompanied by an entropic penalty from restriction of molecular motion. At higher temperatures, rearrangements of the water structure include mutually compensating changes in enthalpy and entropy. Another hypothesis states that enthalpy-entropy compensation is an intrinsic property of weak non-covalent interactions that form a highly cooperative network [36]. Other explanations have been advanced as well, among them the relatively small free energy window in which measurements of K_{eq} can take place [37] resulting in apparent linear temperature plots of ΔG; or highly correlated errors in the determination of ΔH and ΔG [38]. Although the origins of the enthalpy – entropy compensation remains elusive, the biological significance of this phenomenon is unquestionable – it provides the vital mechanism for the living organism to "buffer" solvent perturbations and mutations.

In the temperature range between 0°C and 40°C ΔG of protein – DNA interaction is typically in the range of −20 kJ mol^{-1} to −60 kJ mol^{-1} The corresponding range of equilibrium constants K_{eq} at 298 K is from 1×10^4 to $1x10^{10}$ M^{-1}, rarely stronger than that [39]. There could be a purely biological reason for the many relatively weak protein – DNA complexes. On one hand, binding has to be strong enough to achieve specificity and to compete successfully with other molecules for binding at the particular DNA target sequence. On the other hand, fine regulation of the DNA functions involves simultaneous or consecutive interactions with different proteins, so that binding should be weak enough as to allow dissociation.

Theoretically, the free energy change of a binding reaction (ΔG) can be partitioned into contributions from intramolecular conformational changes, intermolecular interactions at the binding interface, solvent rearrangements, direct macromolecule-water and macromolecule-ion interactions, protonation effects, etc. However, such partitioning of the ΔG (and also of ΔH, and ΔS) is possible only by computation [40, 41].

A variety of different experimental methods can be employed for determination of the ΔG of protein-DNA binding: filter binding assays, fluorescence titration, isothermal titration

calorimetry, electrophoretic mobility gel-shift assays (EMSA), surface plasmon resonance measurements, analytical ultracentrifugation, etc. In practice, filter-binding assays and EMSA are the most popular.

1.2.2 Enthalpy changes

The enthalpy change (ΔH) of binding is experimentally accessible from the temperature dependence of the binding constant by application of the van't Hoff equation:

$$\Delta H = -R\left[\frac{\partial \ln K_{eq}}{\partial\left(\frac{1}{T}\right)}\right]_P \tag{4}$$

In view of the pronounced enthalpy-entropy compensation discussed in the previous section, such estimates of ΔH are often associated with large uncertainty or not possible at all. A direct and precise measurement of ΔH is achieved by calorimetry (see section 1.3).

The magnitude of ΔH provides quantitative estimation of the redistribution of energy levels of the reacting system with reference to the dissociated state. The energy of the noncovalent forces is intrinsically electrostatic but for reasons of convenience noncovalent interactions are classified as hydrophobic contacts, hydrogen bond, salt bridges, long-range electrostatics etc. The change of the energy states upon complex formation is manifested and detected by the absorption or release of heat. The enthalpic effect of an association reaction is the sum of binding per se and possible conformational and/or protonational contributions. DNA binding proteins are usually positively charged at the face with which they contact the DNA polyanion, so that there is a favorable enthalpic term from electrostatic attraction. For the specific complexes, the protein-DNA interface is highly complementary [11, 22, 42] and the close packing [22, 23] results in favorable van der Waals interactions between nonpolar groups, weak polar interactions between the aromatic groups, formation of oriented hydrogen bonds between the polar groups, and salt bridges between fully charged groups. Since the atomic composition, the type of bonds and the packing density of the protein-DNA interface is very similar to those found for globular proteins, by analogy to results obtained with proteins, all these interactions are enthalpically favorable at all temperatures if hydration effects are not considered.

However, dehydration of the protein-DNA interface (release of water molecules that are tightly associated with the macromolecular surface of the free complex partners) is an endothermic process. The magnitude of the enthalpic effect caused by dehydration depends on

the temperature and the polarity of the macromolecule surface being dehydrated. [43]. The net enthalpic effect of binding comprising the formation of intramolecular contacts (exothermic) and water release (endothermic) could be positive or negative due mainly to the temperature variation of the hydratation contribution. A statistical survey on the thermodynamic data collected by ITC for 80 protein – DNA complex entries in the proNIT data base shows an average ΔH of −21 kJ mol^{-1} at 25°C and pH 6 – 8. However, the range of measured ΔH is enormous. For example, binding of λcI repressor to its target duplex was enthalpically favorable with ΔH of -125 kJ mol^{-1} [44] while binding of TBP was strongly opposed by ΔH of 210 kJ mol^{-1} [45]. It should be stressed that with few exceptions, the enthalpy measured by ITC could not be attributed solely to the contacts between protein and DNA groups and the concomitant changes in hydration. Conformational adaptation transitions of the protein and DNA bending contribute to the overall enthalpic term. Moreover, the enthalpic effect of water molecules trapped in cavities at the interface is far from being understood. In the Results section a possible experimental strategy to deconvolute enthalpic contributions stemming from different molecular processes is described.

1.2.3 Entropy changes

The entropy (ΔS) of a binding reaction can be determined as:

$$\Delta S = \frac{\Delta H - \Delta G}{T}$$

if ΔG and ΔH at T are known with accuracy, as it is often the case with ITC experiments. If ΔH is not known, ΔS is calculated from the temperature dependence of ΔG obtained by non-calorimetric methods:

$$\Delta S = -\frac{\partial \Delta G}{\partial T} \qquad (5)$$

As already mentioned, the latter calculation is often flawed due to the relative insensitivity of ΔG to changes in temperature in the physiological temperature range.

The entropy of protein – DNA complex formation accounts for the change of translational, rotational, vibrational and conformational degrees of freedom of all components of the system upon binding, including protein, DNA, water molecules and ions. Eliminating the contribution of the solvent entropy, it is intuitive to assume that the translational and rotational entropy decreases and opposes binding due to the mutual restriction of the degrees of freedom of the macromolecules. The exact magnitude of this term, however, is still much debated. One "classical" view considers the two binding molecules (protein and DNA) to represent two

kinetic units, which form one kinetic unit (complex) upon binding. This estimate, called "cratic" entropy [46], yields a value of −33 J mol^{-1} K^{-1} for ΔS of a binary reaction, which at room temperature (25°C) results in approximately −10 kJ mol^{-1} in terms of TΔS [46]. A much higher value for the rotational/translational penalty of association, 380-420 J mol^{-1} K^{-1} is calculated by application of ideal gas statistics to the macromolecular system [47]. Obviously, the one order of magnitude difference between these estimates leads to very different views about the balance of energy contributions in a binding reaction. Experimental and computational work aimed at solving the dilemma put the decrease of rotational and translational entropy in the range 17-33 J mol^{-1} K^{-1} [48]. As for the vibrational contribution, it might be positive (favorable) or negative (unfavorable), depending on whether binding leads to "loosening" or "stiffening" of the two molecules in the complex, respectively. Tight packing of amino acid side chains and DNA groups at the complex interface causes freezing of rotational bonds and backbone degrees of freedom, and is thus entropically unfavorable. In systems where structural adaptation takes place, there will be an entropic contribution from changes of the conformational entropy, which might be negative or positive.

In solution, there is substantial positive entropic contribution from release of ions and water molecules from the binding interface. It has been found that the entropic effect of hydration of nonpolar groups is negative at biological temperatures and is decreasing in magnitude with the temperature [49]. This effect is attributed to the formation of compressed, low entropy layer of water around the nonpolar chains, which becomes gradually less organized with increasing temperature. In the case of polar groups the entropic effect is also negative, but with a smaller and opposite temperature dependence [49]. The polar groups are highly hydrated thus "fixing" a certain amount of water molecules, which leads to an entropic penalty. The small increase of this penalty with the temperature could be explained with the higher relative restriction of the water molecules, which at higher temperatures are more mobile in the bulk solvent. Altogether dehydration provides the binding reaction with substantial entropic "push" at low temperatures. However, the complementarity of interacting surfaces may not be optimal and dehydration may not be complete. Trapped water molecules will necessarily lower the entropy of binding.

The overall entropic effect is the sum of all mentioned contributions. It varies with the temperature due mainly to the temperature dependence of dehydration. Other less general entropic effects can arise from trapping/expelling of ions and small solutes

1.2.4 Heat capacity changes

The heat capacity change (ΔC_P) is a measure of the temperature dependence of the enthalpy and the entropy changes of binding (eqn 6) at constant pressure.

$$\Delta C_P = C_P^{comp} - (C_P^{pr} + C_P^{DNA}) = [\partial(\Delta H)/\partial T]_P = [T\partial(\Delta S)/\partial T]_P \quad (6)$$

where C_P^{pr}, C_P^{DNA} and C_P^{compl} are the heat capacities of the protein, DNA and their complex respectively. Experimentally, ΔC_p is most reliably calculated from calorimetric data by the Kirchoff's equation:

$$\Delta C_P = \frac{\partial \Delta H}{\partial T} \quad (7)$$

In principle this parameter is accessible also from non-calorimetric binding data by:

$$\Delta C_P = \frac{\partial^2 \Delta G}{\partial T^2} \quad (8)$$

but the precision is very low.

The heat capacity change has attracted special attention because its accurate determination is crucial for a rigorous description of the enthalpy, entropy and free energy changes of the binding reaction. The experimentally obtained values of ΔC_P for protein-DNA association are negative when using the free components as the reference state (eqn 6). Similarly to the other thermodynamic parameters, ΔC_P can also be partitioned into contributions arising from intermolecular interactions in vacuum, conformational changes and hydration effects. The large negative ΔC_P of binding is ascribed mostly to the loss of bound water molecules from the interacting surface areas in the complex [33]. The conformational changes that the protein and the DNA might undergo and the intermolecular interactions (vdW, electrostatic) are considered as smaller contributors and their qualitative (positive or negative) and quantitative effect on ΔC_P is difficult to predict. The magnitude of ΔC_P should therefore roughly correlate with the energetic cost of dehydration of molecular surface [50, 51]. The heat capacity change of dehydration is itself a function of temperature, but this temperature dependence is small and can be neglected in the temperature interval of typical binding experiments. In a wide temperature range the contributions from dehydration of polar, aromatic and aliphatic groups to ΔC_P keep their sign-negative for the nonpolar and positive for the polar groups. The absolute value of the contributions to ΔC_P decrease with increasing the temperature [49].

ΔC_P (and also ΔH and ΔS) can be estimated by semi-empirical methods from the structure of the complex and its components in isolation [50, 51], although such estimates may be ambiguous [38, 52], eqn 41. These correlations are valid only in cases where the binding

macromolecules are considered as rigid bodies and their surfaces are enough geometrically complementary to exclude all the hydration water from the interface. Values of ΔC_P for protein-DNA association tend to be larger than predicted from the burial of polar and nonpolar surface alone [45, 53-56]. Since the conformation of protein and DNA often change when forming a complex [17-20], conformational changes are invoked to explain the observed discrepancies between measured and structure-based estimates of ΔC_P. For example, if binding is coupled to substantial refolding in parts of the protein domain, the heat capacity of refolding (negative) will add to the heat capacity of binding between complementarily shaped molecules. Furthermore, association may change the vibrational modes and the thermal fluctuations of the macromolecule in the complex as compared to the dissociated state. This may also increase or decrease ΔC_P [57, 58] depending on whether the molecules become "loosened" (partially unfolded) or "stiffened" (partially refolded). Contrary to the expectations that the more "free" chains would have bigger C_P, experimental data for protein unfolding [49] shows larger heat capacity for the more compact structures of the folded protein compared to the unfolded structures. This means that the "stiffening" or "packing" of the protein and / or DNA (without the hydration contribution - in vacuum) in the course of the binding would have a positive contribution to ΔC_P and vice versa. Such conformational rearrangement would have much more pronounced effect on the C_P of the binding reaction components in terms of hydration (in the case of "loosening" or partial unfolding) or dehydration (in the case of "stiffening" or partial refolding). These hydration / dehydration effects would be larger and of opposite sign to the effect arising from the conformational changes in the macromolecules alone (in vacuum). Namely a refolding accompanying binding would have net negative contribution to the overall ΔC_P of the reaction and could be subdivided into small positive "stiffening" contribution and large (outside the interface) dehydration contribution. The opposite would be valid in the case for an unfolding accompanying binding. This can be briefly summarized in the following table:

	Contributions to ΔC_P of binding		
	conf. changes (in vacuum)	hydration / dehydration	net
"rigid body" binding		– – –	– – –
"loosening" (partial unfolding)	–	+ +	+
"stiffening" (partial refolding)	+	– –	–

where the number of + or – characters refers to the approximate quantitative effect on the ΔC_P. Vibrational effects are not restricted to protein and DNA. The decrease of translational and soft vibrational modes of water molecules near to or at the binding interface, can be substantial [59]. A recent theoretical study hints at a large heat capacity effect of DNA dehydration [60]. Finally, structural interpretation of heat capacity changes can be severely complicated if the conformation of both protein and DNA are temperature-sensitive. Many DNA-binding proteins are flexible and sometimes partly or even completely unfolded in isolation and have low thermal stability [19].

1.3 Calorimetric methods and their application for studying protein–DNA binding reactions

Analysis of the energetics of protein-DNA association requires the knowledge of all the conformational transitions of the complex and its free components in as large a temperature range as possible. In particular, knowing the heat capacities of the complex and its components is necessary to interpret discrepancies between measured and calculated heat capacity changes [5, 54-56] . The best way to thermodynamically "decipher" the complicated network of interactions and forces of the protein – DNA binding is to complement the isothermal titration calorimetry (ITC) with differential scanning calorimetry (DSC) and other "thermal melting" methods. Only from the combined results of such studies performed on the complex and its components can one reliably correlate energetics with structure.

1.3.1 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is the most powerful technique for studies of the thermodynamics of bimolecular interactions. It measures the absorption or release of heat when two molecules bind to each other. Heat change is a universal feature of chemical processes and, as experimental signal, it allows characterization and quantification of a binding event in solution without the need of tagging, chemically modifying or immobilizing the molecule.

The principle of the experiment is rather simple. A solution of one of the interacting molecules (called "ligand", L, for convenience) is added at constant temperature with a precision syringe into a cell containing a solution of the other molecule (called "macromolecule", M). A computer controls the volume and duration of each injection. The solution in the sample cell is constantly stirred by the rotating syringe. The buffer in which the two components are dissolved should be the same so that the heat effect of mixing of the two solvents is zero. Apart from the

sample cell there is a reference cell of equal volume filled with the same buffer as the sample cell. No additions are made to the reference cell; it serves as a heat capacity reference only. Both cells are mounted in an adiabatic chamber, whose temperature is kept constant. When the ligand binds to the macromolecule heat is released or absorbed in direct proportion to the amount of complex formed. The observable signal of the instrument is the electrical power (in units of Watts, J s^{-1}) required to keep a small temperature difference between the sample cell and the reference cell constant, $\Delta T = T_S - T_R = const$. While no reaction is in progress $\Delta T = const$ and, therefore, the electrical power is constant over time. The heat generated after mixing displaces ΔT from its equilibrium value. Electrical power is fed to the heaters of the two cells in order to re-establish the magnitude of ΔT at thermal equilibrium. Figure 1a shows a representative titration experiment. The area of each peak corresponds to the power signal integrated over the time from the moment of injection to reaching of equilibrium (usually a few minutes) and equals the heat (q, in units of J or cal) associated with the addition of each aliquote of the ligand. Integrated heats are shown in Figure 1b.

In the course of the titration, as the ligand concentration in the cell increases, the macromolecule becomes progressively saturated. The heat signal diminishes until only the background heat of dilution and other non-specific effects are observed (Figure 1); at this point there are virtually no free binding sites available for binding. Usually the titration proceeds until the macromolecule in the cell is over-saturated and there is no more heat from the binding process. As with other experimental signals that are used to evaluate the progress of binding, the amount of heat after each injection is also a valid molecular “marker”. Plotting the heat effect of each injection vs the respective ratio “ligand” / ”receptor” results in a sigmoidal curve with inflection at the point where the amounts of L and M in the cell are equal (in the case of one to one binding) (Figure 1b). Analysis of this isotherm [61, 62] by standard least-squares non-linear regression methods provides the equilibrium constant of the binding reaction (K_{eq}), its stoichiometry (n) and the molar enthalpy (ΔH) (eqs 9,10):

$$\frac{dq}{dL_T} = \Delta HV\left[\frac{1}{2} + \frac{1 - X_R - r}{2\sqrt{(X_R + r + 1)^2 - 4X_R}}\right] \tag{9}$$

where dq is the heat effect of a single injection; dL_T the total accumulated molar concentration of ligand in the cell; V the current volume of solution in the cell; $X_R = [L_T]/[M_T]$; M_T the total molar concentration of macromolecule in the cell;

$$r = 1/K_{eq}[M_T]; L_T = L + nM_T\left(\frac{K_{eq}L}{1 + K_{eq}L}\right) \tag{10}$$

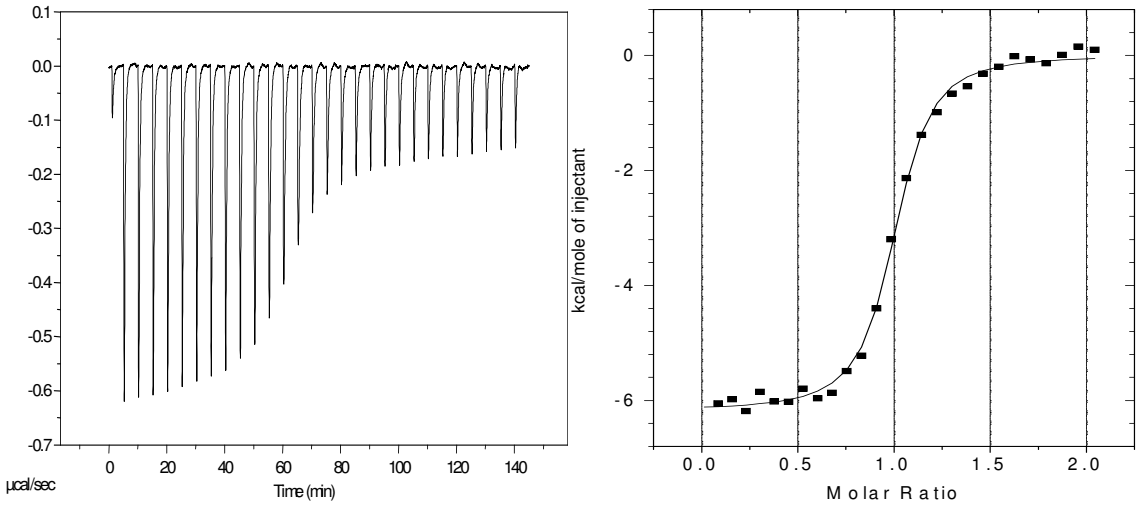


Figure 1. Representative ITC data. 29 injections of ligand solution (L) are added to macromolecule (M) solution in the ITC cell. The area underneath each injection peak (Fig.1a-left panel) is equal to the total heat released for that injection. When this integrated heat is plotted against the molar ratio of ligand added to macromolecule in the cell, a complete binding isotherm for the interaction is obtained (Fig. 1b-right panel). The one site model was used to fit the data. The solid line is the calculated curve using the best-fit parameters.

Accurate determination of the enthalpy (ΔH), equilibrium constant (K_{eq}) and stoichiometry (n) is readily providing values also for the free energy (ΔG) (eqn 1) and entropy (ΔS) (eqn 2), thereby completing the thermodynamic profile of the molecular interaction at the given temperature in a single experiment. For such full thermodynamic description in one experiment the binding system should fulfil certain requirements. Very tight binding reactions (K_{eq} higher than 10^9 M^{-1}) exhibit very steep binding isotherms, out of which no accurate value for the binding constant can be extracted, but accurate estimate of the enthalpy change is possible. One way to overcome this problem is to reduce the concentrations of the macromolecule and the ligand, but this may decrease the observable heat effect to values below the sensitivity of the instrument. Very high affinity binding can be characterized by displacement experiments [63, 64], in which the strongly binding ligand is titrated into a solution containing the receptor already saturated with a weaker ligand. Given that binding of both ligands exhibits a large enough enthalpy change the binding constants of both ligands can be obtained. At the other extreme, very weak binding results in a very shallow binding isotherm, not reaching saturation; estimates for both enthalpy and association constant are inaccurate. In this case increasing the concentrations of the reacting species is required, but aggregation, high viscosity and low solubility are possible limiting factors. As an empirical rule, good estimates of

the enthalpy change and the binding constant are obtained for: cK_{eq} = 10 to 100, where "c" is the total concentration of binding sites in the cell and K_{eq} is the equilibrium constant.

When the ITC experiment is repeated at different temperatures, the heat capacity change accompanying binding can be calculated by Kirchoff's equation (7).

1.3.2 Differential scanning calorimetry

Calorimetric methods are the only means for direct measurement of the enthalpy change (ΔH) of biomolecular reactions. While ITC is used primarily for determination of thermodynamic binding parameters, differential scanning calorimetry (DSC) directly measures the thermodynamics of conformational transitions of macromolecules [65, 66]. In the context of this work we have investigated the unfolding of protein, DNA and their 1:1 complex by DSC.

As in the case of ITC the scanning calorimeter has two cells with equal volumes – a sample cell and a reference cell. During the experiment both cells are subjected to strictly controlled quasi-adiabatic heating with a specified rate, usually 0.5-1.5 K min^{-1}. The observable signal is the electrical power difference (in units of Watts) between the heaters of the cell containing the dissolved macromolecule and the heaters of the reference cell (filled with buffer only). Since the rate of increasing the power is predefined, this differential mode of measuring results in correlating the power difference (ΔP) with the heat effect due solely to the unfolding of the macromolecule. The electrical power is the first derivative of the energy (Q) over time (t) (eqn 11a); thus in an experiment where the heating rate (K min^{-1}) is known and given that the enthalpy change (ΔH) equals the heat transfer (ΔQ) at constant pressure one can directly relate the power to the heat capacity of the solution (eqn 11).

$$P = \frac{\partial Q}{\partial t} \qquad C_P = \frac{\partial H}{\partial T} \qquad dC_P = \frac{dP}{(dT/dt)} \tag{11}$$

The differential scanning calorimetry measures the heat capacity of the molecule under study as a function of temperature. The parameter of interest is $C_P(T)$, the partial specific heat capacity, J K^{-1} g^{-1}, of the solute. Since the raw signal is the sum of all the contributions to the heat capacity one has to subtract the heat capacity of the solvent. For this purpose, knowledge of the partial specific volumes of the macromolecule and the solvent as well as their temperature dependencies is required.

The value of $C_P(T)$ is an indicator of the conformational state of a protein in the temperature interval of the experiment. The partial specific heat capacity of the native protein is

lower than that of the denatured protein. Precise measurements of the absolute value of the heat capacity for numerous single domain globular proteins showed that the C_P(25°C) varies in the range 1.2–2.3 J K^{-1} g^{-1} and increases linearly with temperature with a slope of (6–8)x10^{-3} J K^{-2} g^{-1} [49, 67]. On the other hand C_P(T) of the unfolded polypeptide chain can be calculated from the amino acid sequence by summing up the heat capacities contributions of the amino acid residues and the peptide bonds [68, 69]. Having on hand an accurate value for at the post-transition temperatures one can compare it with C_P calculated from the amino-acid composition. Experimentally if the measured C_P(T) is lower than the calculated C_P(T), the denatured protein may have some residual structure and is not completely unfolded.

Figure 2 shows a typical DSC curve. It can be divided in three parts: pre-transition, transition and post-transition. The pre-transition part represents the heat capacity function of the

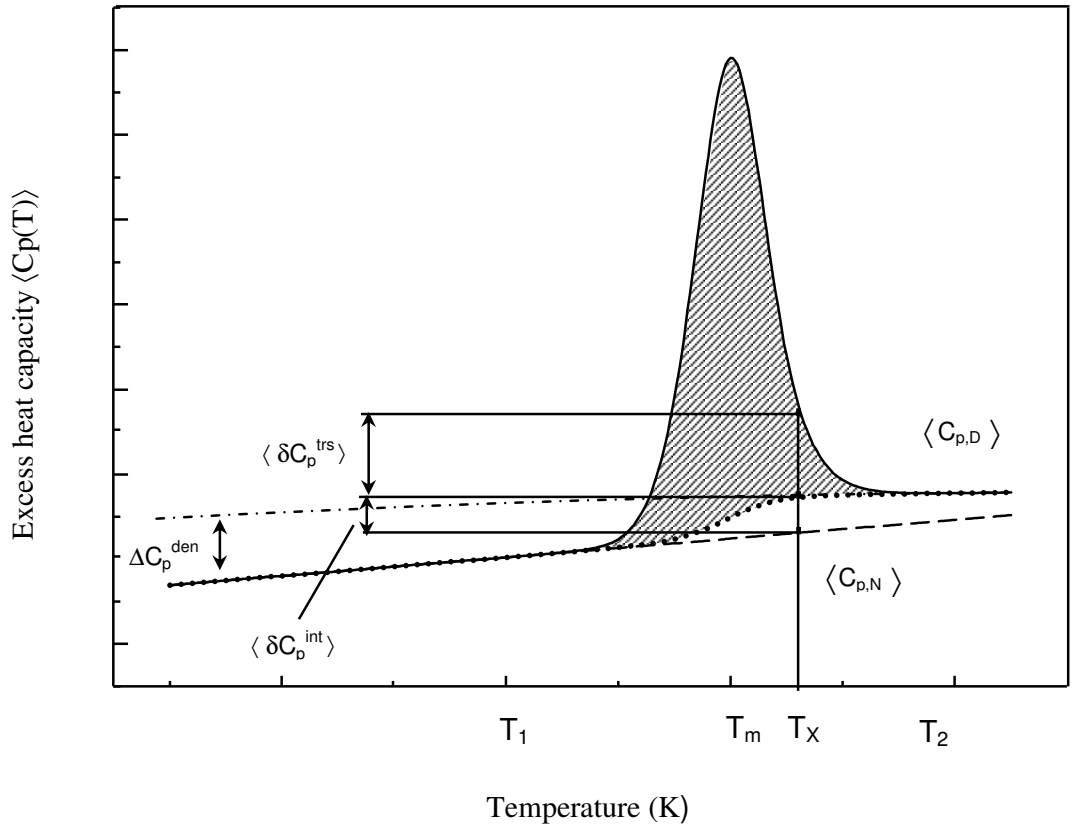


Figure 2. DSC curve of protein thermal denaturation. The heat capacity of the native state, $\langle C_{p,N}(T)\rangle$, shows a linear temperature dependence (dashed line). The heat capacity of the denatured state, $\langle C_{p,D}(T)\rangle$, is approximated by a quadratic function (dashed-dotted line). The functions $\langle C_{p,N}(T)\rangle$ and $\langle C_{p,D}(T)\rangle$ can be extrapolated into the transition zone between T_1 and T_2 in proportion to the progress of transition (dotted line). The excess partial heat capacity $\langle C_p(T)\rangle$ is composed of the intrinsic excess heat capacity $\langle \delta C_p^{int}\rangle$, which accounts for all the molecular species that become populated in the progress of transition, and from the transition excess heat capacity, $\langle \delta C_p^{trs}\rangle$, originating from the increased enthalpic fluctuations of the system when the protein changes between different conformational states in the course of thermal denaturation. The relative magnitudes of $\langle \delta C_p^{int}\rangle$ and $\langle \delta C_p^{trs}\rangle$ at temperature T_x are indicated (Reproduced from [4]).

folded protein, $C_{p,N}$. As mentioned earlier it has some characteristic slope. The post-transition part represents the heat capacity of the unfolded state, $C_{p,D}$. The absolute values of the pre- and post-transitional functions are the summed contributions to the heat capacity of atoms and covalent bonds, intramolecular noncovalent interactions, macromolecular thermal motions and vibrations, and the heat capacity due to hydration of the protein surface. As shown in Figure 2 $C_{p,N}$ and $C_{p,D}$ differ numerically. The reason is that in the unfolded state the intramolecular interactions are virtually zero, the polypeptide chain motions are much higher and more molecular surface is exposed to water. The heat capacity increment accompanying protein denaturation, $\Delta Cp = C_{p,D} - C_{p,N}$ is positive. It is typically in the range 0.3 to 0.7 J K^{-1} g^{-1}. The increase of heat capacity upon unfolding is dominated by hydration effects.

Experimentally, ΔC_p is determined by extrapolation of $C_{p,N}$ and $C_{p,D}$ into the transition zone (around T_m) and taking the difference $C_{p,D} - C_{p,N}$. The partial molar heat capacity of the native protein can be approximated by a linear function (dashed line), and that of the denatured protein by a quadratic function (dash-dotted) [70-72], its curvature can be calculated from the amino acid composition assuming that the unfolded state is fully hydrated [73].

The bell shaped transition part of the heat capacity function can be divided in two components. The first component accounts for the distribution of the molecular species that become populated in the progress of the transition from the folded to the unfolded state. It is simply the increase of the heat capacity from the value for the folded state to the value for the unfolded state without exceeding the latter. This component is usually referred to as the intrinsic excess heat capacity. The second contribution to $C_P(T)$ originates from the increased fluctuation of the system when the protein changes between different enthalpic states in the course of thermal denaturation. It represents the "burst" cooperative uptake of heat with a relatively small change in the temperature needed for the destruction of the intramolecular interactions. This component is called transition excess heat capacity and is usually much larger than the intrinsic excess heat capacity as indicated in Figure 2.

The excess enthalpy of the thermal denaturation corresponds to the area of the peak above the line of the intrinsic excess heat capacity in Figure 2. The numerical value of this enthalpy is the 'true' calorimetric estimate of the denaturation enthalpy and is model independent, i.e. it does not depend on the shape of the transition curve. Mathematically it is obtained by integration of the heat capacity temperature function.

$$\langle \Delta C_P(T) \rangle = C_P(T) - C_{P,0}(T) \qquad \langle \Delta S(T) \rangle = \int_0^T \langle \Delta C_P(T) \rangle d \ln T \tag{12}$$

Where $\langle \Delta C_P(T) \rangle$ stands for the excess heat capacity, $C_P(T)$, relative to a state (intrinsic excess C_P) considered as reference, $C_{P,0}(T)$. This integral can also be analyzed according to the van't Hoff equation yielding another model-dependent value of the transition enthalpy:

$$\Delta H_{vH} = 2T_M \sqrt{R(\langle \Delta C_P(T_M) \rangle - 0.5\Delta C_P)} \quad (13)$$

where T_M is the temperature at the maximum of the heat capacity curve and R is the universal gas constant. The van't Hoff model (eqn 13) assumes equilibrium between only two forms: folded and unfolded protein. Comparing the values of the calorimetric and van't Hoff enthalpies provides us with information about the cooperativity of the transition. If they are equal the unfolding is considered to be a highly cooperative two state process; if the van't Hoff enthalpy is higher than the calorimetric enthalpy, this may indicate an irreversible process taking place; if the van't Hoff enthalpy is smaller than the calorimetric enthalpy, unfolding is not strictly two state process and some intermediate states become significantly populated during the transition.

Integration of the excess heat capacity function with respect to lnT yields the calorimetric entropy of unfolding at T_M: $\Delta S = \int_0^T \langle \Delta C_P(T) \rangle d \ln T$

In this way, since ΔG at T_M and ΔC_p are known, the thermodynamic characterization of the unfolding transition is complete.

The same calculations can be performed with the heat capacity function obtained for DNA melting. It should be noted that melting of DNA includes simultaneous melting of the double helix and strand dissociation. The peculiarity of bimolecular unfolding (or higher order unfolding) is that the heat capacity peak is not symmetrical, the maximum of the excess heat capacity function appears at temperatures slightly higher than the temperature where the system is half unfolded, and T_M depends on the concentration [74]. Integration of C_p over T or lnT yields the calorimetric values of ΔH and ΔS, respectively. However, calculation of the model-dependent enthalpy, entropy and free energy requires the bimolecular order of the process to be considered explicitly.

In spite that DSC is used mainly to quantify the conformational stability of macromolecules, the method has proved a valuable tool to analyze biomolecular interactions as well [4, 75]. The heat capacity function of the protein-DNA complex is not a simple geometrical sum of the C_p traces of the two components. If the mutual stabilization of the complex partners is high, the dissociation of the complex is tightly coupled to the unfolding of its components resulting in a single peak in the heat capacity trace at a temperature above the denaturation temperatures of the protein and DNA in isolation. Careful deconvolution of the

observed thermograms gives insight in the thermodynamic forces that drive the association reaction and can provide structural and mechanistic information about the interacting molecular species [76-78]. In a simplistic treatment, inegration of the melting trace results in the enthalpy of the protein and DNA unfolding / dissociation at the temperature of the heat capacity peak. If the enthalpies of protein and DNA melting (after extrapolation to the melting temperature of the complex) are subtracted, one should obtain the enthalpy of the association reaction alone. However, in the majority of systems studied so far, the DSC trace has a complicated shape and exhibits several overlapping transitions. Moreover, even if the dissociation and the denaturation take place in a narrow temperature range, there is no a priori information about a possible redistribution of sub-states in the melting zone.

1.4 Kinetics of the protein – DNA reactions

The origins of the stability and specificity of the protein-DNA complexes, as well as the mechanism of their binding can never be fully understood without a combined study of the structure, energetics and kinetics of the complex formation. While the energetics provides information about the stability of the complex and whether binding is linked to enthalpic and entropic cost or gain, kinetic studies reveal the rates of the association and dissociation reactions comprising the equilibrium. Since kinetics deals with the dynamics of the complex, it could shed some light also on the mechanism of the reaction. In the case of a bimolecular reaction we have:

$$P + D \underset{k_1}{\overset{k_{-1}}{\leftrightarrows}} PD \qquad K_{eq} = \frac{k_1}{k_{-1}} = \frac{[PD]}{[P][D]} \tag{14}$$

The kinetic constants of association, k_1, and dissociation, k_{-1}, represent the rates at which the concentration of a given reactant is decreasing or increasing to reach the equilibrium concentration:

$$\frac{d[P]}{dt} = -k_1[P][D] + k_{-1}[PD] \tag{15}$$

where [P], [D] and [PD] are the concentrations of free protein, free DNA and complex, respectively. The rate of encounter, k_1, of the two components in the ideal case of diffusion limited reaction can be calculated in terms of the Smoluchowski theory [79]. In the diffusion limit, every single encounter is productive, i.e. binding takes place whenever the two molecules collide driven by diffusion. The diffusion limited rate constant of association (k_e) depends only on the effective radii of the molecules, r_A and r_B, and their diffusion coefficients, D_A and D_B:

$$k_e = \frac{4N}{1000}(D_A + D_B)(r_A + r_B)\pi \quad (16)$$

where N is the Avogadro number. For typical single domain proteins and short duplexes we have *r* values around 10 – 20 nm and diffusion coefficients at room temperature around 10^{-6} cm^2 s^{-1} which yields a value for k_e of around 10^9 M^{-1} s^{-1}. In reality, this number is usually lower. First, only specific "patches" on the surface of both macromolecules can interact with each other, and therefore, dissociation/re-association or rearrangement of weak initial encounter complex are necessary. Second, only certain conformations of the protein and/or DNA may be "binding-competent". They are in equilibrium with non-competent conformations and the effective concentration of the "competent" form may be low. Accordingly, conformational selection [80-82] takes place. However, association can be greatly accelerated, to the diffusion limit or beyond, because of the following. First, in the living cell the translational movements of proteins and DNA are spatially restricted within a specific compartment-nucleus or mitochondria and the effective concentrations can be very high [83]. Second, long-range electrostatic attraction between oppositely charged molecules could "funnel" the reactants together resulting in an "oriented diffusion". Third, studies on specific and non-specific protein – DNA binding reactions have indicated that the initial contact represents an energy minimum on the reaction coordinate. In the most probable case, initial encounters have very low association equilibrium constants K_A due solely to high k_{-1} while association rate constants k_1 are similar. This results in a fast dissociation after the encounter and another "try" in the vicinity of the first contact; in other words the protein is "sliding" along the DNA chain [84] until it finds the sequence with the deepest energy minimum (and the lowest k_{-1}), namely its cognate site. It has been proposed that DNA-binding proteins can slide along the DNA after an initial encounter driven purely by electrostatics and in this way the diffusion toward the target site becomes one-dimensional [84].

Unlike k_1, the dissociation rate constant, k_{-1}, depends mainly on close range interactions between the two molecules. It has been shown that k_{-1} is intimately related to the specificity of the complex [85-87].

Fast kinetics studies require special instruments and methods. Presently, the most common approaches involve stopped-flow mixing measurements. In such an experiment, the two binding species are placed in separate syringes. They are mixed rapidly (under pressure) in a mixing chamber. The observable signal is the change of fluorescence, ellipticity etc., accompanying binding or folding/unfolding. Usually reactions are taking place in the time interval of milliseconds to minutes, but there are extremely fast or slow reactions reaching

equilibrium in the order of microseconds or hours, respectively. The time limit of the stopped-flow instruments is the “dead time”; this is the time between the actual start of the reaction and the start of a reproducible change of signal-usually around 1 ms. Before this time the signal is “unstable” as the mixing of the solutions is not complete. The observed signal change is due to the change in the surroundings of the molecular marker (intrinsic fluorophore for example) upon binding. Therefore, for the reaction to be observable the molecular marker should be sensitive to the process of binding or folding/unfolding. In the case of fluorescence stopped-flow, intrinsic fluorophores of the protein (tryptophan and tyrosine residues) serve as markers. External fluorophores (special fluorescent dyes) can be attached to one or both of the molecules. After mixing, the signal decreases or increases as the system approaches equilibrium. The resulting reaction trace can be described by an exponential function or by the sum of exponential functions.

2 BACKGROUND AND OBJECTIVES OF THE PRESENT STUDY

2.1 Biological role of the Tn916 integrase

Conjugative transposons [88-96] are widely distributed in nature and are responsible for spreading antibiotic resistance among various pathogenic bacteria. During transposition these DNA elements can move (excise and insert) from one place to another within a single genome as well as be transferred to other cells of the same or different species. The Tn916 transposon is 18 kilobases in length and carries resistance to tetracycline [96]. Tn916 integrase (transposase) that performs the excision and joining reactions of the Tn916 transposon is a two-domain protein. The catalytic domain is positioned in the C-terminal part of the molecule and recognizes and cleaves inverted repeats at the junction between the transposon and the rest of the bacterial chromosome. The DNA-binding domain (Int-DBD) consists of the N-terminal 74 residues and specifically recognizes repeated 11 bp DNA stretches in the transposon arm located 90 and 150 bp from the left and the right of the transposon, respectively.

2.2 Structural features of Int-DBD and its complex with DNA

Structurally, the most interesting feature of Int-DBD is its unusual mode of recognizing DNA and its close similarity to a common RNA binding motif which recognizes A-shaped RNA [97-100]. Unlike most small DNA binding domains, which utilize α-helices as the principal binding motifs, Int-DBD inserts a three stranded antiparallel β-sheet into the major groove of a B-form DNA. This type of recognition constitutes a novel binding motif which is encountered only in two other proteins [101, 102].

The structures of the free protein and its complex with a 13 bp long DNA duplex have been solved by NMR [1]. The physical size of the Int-DBD is 41Å×23Å×34Å. An antiparallel β-sheet packs against two structurally ordered loops and an α-helix. The protein adopts Lβββlα topology, where L stands for loop [2]. The exterior face of the β-sheet forms a concave surface with turns T1 and T2 protruding into the solvent (Figure 3).

The strands of the β-sheet are formed from residues Glu 17–Gln 19 (strand β1), Tyr 25–Ile 30 (strand β2) and Pro 36–Ser 41 (strand β3). The connecting turns comprise residues Arg 20–Arg 24 (turn T1) and Asp 31–Glu 35 (turn T2).

The sheet is preceded by the N-terminal loop L1–(residues Lys 4–Gly 16). Another extended loop (loop L2, residues Trp 42–Ser 59) projects away from the body of the protein. Loop L2 connects strand β3 to a C-terminal helix (residues Leu 60–Gln 68).

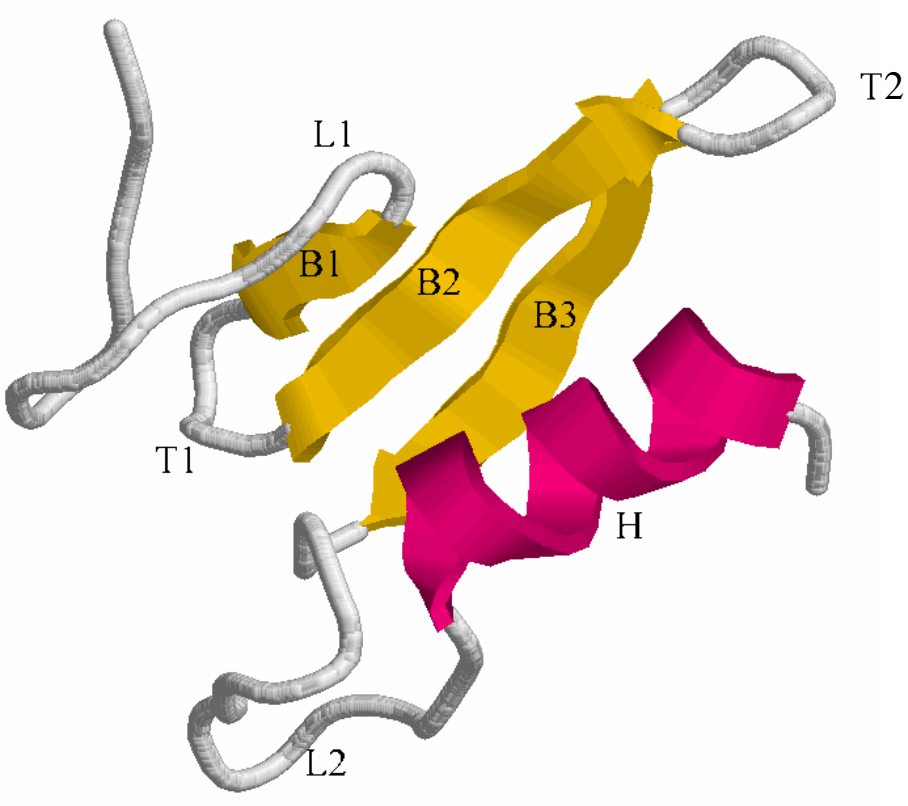


Figure 3. Schematic diagram of the structure of Tn916 integrase N-terminal domain (Int-DBD). Color coding: C-terminal α-helix (H), red; three stranded β-sheet (strands B1, B2, B3), yellow; loops and turns (L1, L2 and T1,T2 respectively), light gray. Picture was created by RasMol 2.7.2.1 using the coordinates deposited in PDB, accsess code 2bb8 [2].

Int-DBD binds to one side of the duplex, burying approximately 2300 $Å^2$ of solvent-accessible surface area. The β-sheet is positioned within the major groove formed by base pairs 3–6. The DNA duplex is bent by ~35°, enabling additional protein–DNA contacts at the 3' side of the duplex from residues in the extended ordered loop L2. With the exception of the helix and turn T2, all other secondary structural elements contact DNA. The overall structure of the complex is presented in Figure 4.

Residues located in strands β2, β3 and in turn T1 make sequence specific contacts with DNA bases and form a concave-shaped surface that fits in the major groove [1]. There are three base-specific protein–DNA hydrogen bonds in this interface. The hydroxyl group of Tyr 40 is positioned as to accept a hydrogen bond from the N4 amino group of C21. A second hydrogen bond is made by the amino group of Lys 28, which is close to the O6 and N7 atoms of G3. A

third hydrogen bond is likely made by the side chain of Arg 20. This residue is inserted into the major groove. In several of the NMR conformers the guanidine group of Arg 20 donates hydrogen bonds to the O6 and N7 atoms of G6, while in others, it is hydrogen-bonded to the O4 atom of T7. Only one protein–DNA interaction is made to the complementary DNA strand: the main-chain amide of Phe 38 contacts the DNA by donating a hydrogen bond to the phosphate of C21.

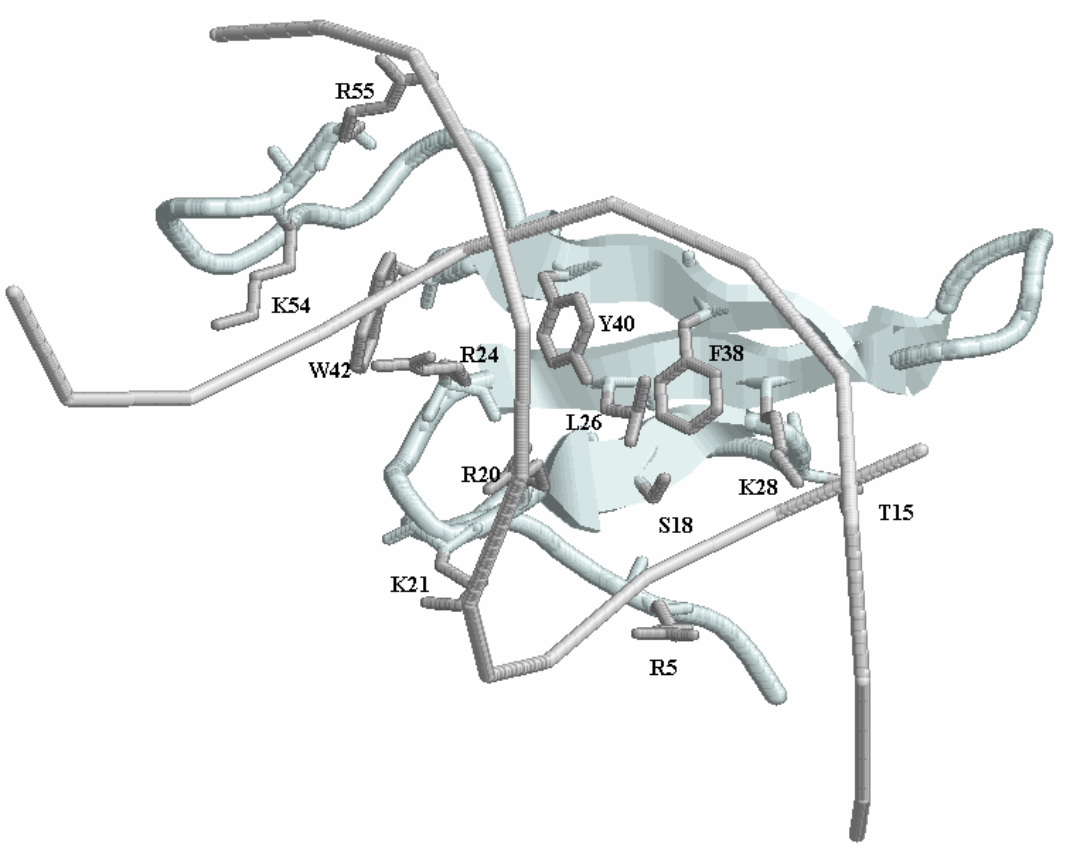


Figure 4. Schematic diagram of the structure of the Int-DBD-DNA complex with a view of the DNA contacting amino acid side-chains. Color coding: Protein backbone without the C-terminal α-helix, light cyan; DNA duplex backbone, light gray; side-chains of the residues contacting DNA with single letter amino acid coding and numbering, dark grey. Picture was created by RasMol 2.7.2.1 using the coordinates deposited in PDB, accsess code 1b69 [1].

Contacts between the DNA and the side chains of Lys 54 and Arg 55 (loop L2) are facilitated by bending at the T7/A8 base. This distortion is stabilized also by the side chain of Arg 24 from turn T1, which interacts with the phosphates of T18 and T19. Sequence-dependent bending might be crucial for high-affinity binding since the T7/A8 base step is fully conserved in all of the Int-DBD binding sites.

Comparison of the free and DNA-bound protein structures shows some slight change in the structure of the Int-DBD upon binding – a displacement of turn T2, which is pulled away from the C-terminal α-helix by ~4.5 Å at its tip. This translation causes the structure of the protein to become more disordered in its bound state – the edge of the hydrophobic core is getting disrupted, destabilizing the C-terminal end of the α-helix, which is two residues shorter in length. Turn T1 at the interface is also pulled toward the DNA by ~1.8 Å, enabling tighter binding.

2.3 Previous biophysical characterization of Int-DBD binding to DNA

Biophysical characterization of the Int916-DBD-DNA complex is needed in order to gain insight into the exact mechanism and affinity determinants that govern the association of the three stranded β-sheet motif with DNA. From equilibrium binding studies [103] the equilibrium dissociation constant (K_D) for the Int-DBD-DNA complex is 40-150 nM, depending on the method used. Careful examination of the NMR structural data singled out 13 amino acids residues that are positioned as to contact the DNA strands. The energetic significance in terms of ΔG of all these local interactions was probed by alanine substitutions (see Table 1 in [103]). The residues span the entire complex interface. Formation of three base-specific hydrogen bonds (Arg 20, Lys 28 and Tyr 40) and two base-specific van der Waals contacts (Leu 26 and Phe 38) is supposed to play the major role of determining the specificity of the complex. The rest of the contacts are predominantly non-specific hydrogen bonds with the main chain DNA phosphate groups (Arg 5, Ser18, Lys 21, Arg 24, Trp 42, Lys 54 and Arg 55) and one deoxyribose van der Waals contact (Thr 15). Combination of the results from the mutagenesis data with structural information yielded a comprehensive description of the Int-DBD-DNA interactions. It appears that the interfacial hydrogen bonds particularly between the β-sheet strands β2 and β3 and DNA, are most important for high affinity binding while hydrophobic interactions are less important. These observations were confirmed by experiments performed with wild type protein and mutated DNA duplexes. The two base-specific hydrogen bonds originating from Lys 28 (strand β2) and Tyr 40 (strand β3) appear the energetically most important since their removal results in a 11 and 14 fold increase of K_D, respectively. Some contacts which are away from the β-sheet-DNA interface also contribute strongly to the overall affinity. Mutations of one residue from turn T1 (Arg 24) and one from the extended loop L2 (Lys 54) are significantly affecting binding – these contacts are considered important for positioning of the C-terminal side of the β-sheet and are facilitated by the DNA distortion. The

contacts from the other side of the interface (with loop L1) seem to be less important than those from loop L2 even though there is some modest penalty for substituting Arg 5 by Ala.

2.4 Objectives of the present study

Here we use the N – terminal domain of the bacterial integrase Tn916 and its target DNA as a relatively simple system for the comprehensive thermodynamic analysis of a protein-DNA complex. With the available structural data for the isolated protein and its complex with a short DNA duplex containing the cognate site, this complex represents and ideal system for investigating general structure-energetics relationships in site-specific protein-DNA recognition. Specifically, this work intends:

- to achieve a complete thermodynamic description of the stability of the complex, including quantification of the enthalpy, entropy and heat capacity changes accompanying binding;
- to ratioanalize the magnitude of the observed enthalpy and heat capacity of binding in terms of the structural features of the complex;
- to investigate the salt dependence of the free energy of binding and its enthalpic and entropic components;
- to measure by calorimetry the thermodynamic signature of mutations eliminating protein-DNA contacts that stabilize the complex;
- to characterize the dynamic properties of the complex by investigating the association and dissociation rates of the wild-type complex and selected mutants;
- to give a detailed biophysical description of Int-DBD in terms of its thermodynamic stability and folding/unfolding kinetics.

3 MATERIALS AND METHODS

3.1 Materials

Most experiments were conducted in standard buffer composed of 50 mM Na-phosphate, 100 mM NaCl, pH 6. The salt dependency experiments were conducted with varying the concentrations of Na^+ in the range 120 mM to 320 mM, while the phosphate concentration was kept 80 mM. For the osmotic stress measurements glycerol was added to the original buffer (of 50 mM Na-phosphate, 100 mM NaCl, pH 6) with concentrations 16% and 28% (weight percent) – the ionic strength was kept constant. Some ITC measurements were performed in 64 mM HEPES (or ACES), 100 mM NaCl, pH 6.0. The pH of samples containing urea or GdmCl was adjusted after addition of the denaturant. Urea and GdmCl concentrations were determined by refractometry. All chemicals were of analytical grade and were used without further purification.

3.1.1 Overexpression and purification of INT-DBD

The sub-cloning of the N-terminal domain of Tn916 integrase (INT-DBD) comprising residues 2-74 and the Cys57Ala mutation has been described elsewhere [1]. *E. coli* cultures were grown for 4 h at 37 °C and induced by addition of 1mM isopropylthio-β-D-galactoside. Cells were harvested by centrifugation, resuspended in 100 mM Tris-HCl (pH 7.2), 5 mM EDTA, 2.5 mM dithiothreitol (DTT), 5 mM benzamidine, and lysed by sonification. The lysate was centrifuged at 15 000 rpm for 1 h and the supernatant containing INT-DBD was loaded on a heparin-Sepharose CL-6B column (Pharmacia; 1.6 cm x 14 cm) pre-equilibrated with 100 mM Tris-HCl (pH 7.0), 5 mM EDTA, 2.5 mM DTT. After thorough washing, elution was achieved with 1 M NaCl in the same buffer. The eluate was dialyzed overnight against 50 mM Tris-HCl (pH 7.0), applied to a LKB-SP-5PW column and eluted with a gradient of 1 M NaCl (0–100%) in 50 mM Tris-HCl (pH 7.0). INT-DBD was purified to homogeneity by reversed-phase HPLC on a Nucleosil 300-5 C8 column (Machery & Nagel) and binary gradients of acetonitrile/H_2O containing 0.1% trifluoroacetic acid. The protein was lyophilized and stored as a powder. Ion spray mass spectrometry showed a single peak of 8555 Da corresponding to residues 2-74 (C57A) and the respective molecular masses for the mutant proteins. Protein concentration was measured by UV absorption at 280 nm in 6M GdmCl using the calculated ε_{280} of 10.81 M^{-1} cm^{-1} or the calculated ε_{280} of the respective mutant.

3.1.2 Mutagenesis of INT-DBD

The plasmid was extracted from the cells with QIAfilter Plasmid Kit and the subsequent mutagenesis was performed with Stratagene QuikChange Site-directed Mutagenesis Kit; using the manufacturer's protocol without modifications. The primers for the mutants were purchased from Microsynth and were used without further purification. Mutations were verified by sequencing in the same company. The overexpression and purification procedures for the mutated proteins were the same as described above.

3.2 Methods

3.2.1 Preparation of DNA duplex and protein-DNA Complex

Single-stranded oligonucleotides were purchased from Metabion GmbH (Martinsried) and were purified by HPLC. For duplex preparation, equimolar amounts of the two complementary strands were mixed and annealed by heating to 70 °C and slowly cooling to room temperature. Concentrations were determined, after complete digestion by phosphodiesterase I, from light absorption at 260 nm (Sigma). The complex was formed by adding the protein to a solution of DNA until the calculated 1:1 ratio was reached.

3.2.2 Circular dichroism measurements

CD measurements were carried out on a Jasco J-715 instrument. Spectra shown are the average of three scans recorded at 5 nm min^{-1}. Thermal unfolding curves were measured by following the ellipticity change upon continuous heating or cooling between 3 and 65–85 °C at a scan rate of 0.5 or 1 °C min^{-1} and with data collection every 20 s. Reversibility of unfolding was checked by repeated scans and was always better than 95%. Thermal melting curves were analyzed as described [104].

In thermal unfolding experiments by CD the continuous heating rate was 0.5 or 1 deg min^{-1}.

3.2.3 Fluorescence spectroscopy

3.2.3.1 Equilibrium measurements

Fluorescence measurements were made on a Perkin Elmer LS 50B luminescence spectrometer. Excitation was at 295 nm and emission spectra were recorded from 300 to 450 nm at 1 nm intervals with 1.2 s integration time. For binding experiments, 2-5 μM protein placed in

the thermostatted cuvette was titrated with 2 μL aliquots of concentrated DNA with continuous stirring. Final dilution was 4% or less. In thermal unfolding experiments by fluorescence the temperature of the Peltier-thermostatted cuvette holder was increased step-wise with 5 min equilibration time between steps. The temperature was controlled by a sensor in physical contact with the sample solution. Denaturant induced conformational changes were monitored from the position of the fluorescence emission maximum induced by excitation at 295 nm, after 2-5 μM protein had been incubated for 12–15 hours at the desired denaturant concentration.

3.2.3.2 Kinetic (stopped – flow) measurements

Rates of protein folding and unfolding as wells as of protein – DNA association were studied at 7 °C with the SX.18MV-R stopped-flow spectrometer from Applied Photophysics (Surrey, U.K.). For the protein folding and unfolding experiments the protein and urea solutions were mixed in ratio 1:5. Twenty different urea solutions were used as the final urea concentration increased from 0.63M to 7.3M. The final protein concentration was 1mM. For the protein – DNA binding experiments equal volumes of equally concentrated protein and DNA solutions were mixed (dead time of <3 ms). For all experiments the change in fluorescence emission above 320 nm was measured upon excitation at 280 nm. Five or more firings were averaged for each kinetic trace. Reported values for the kinetic constants represent mean from experiments done with 3-4 different concentrations of the complex in the range 0.5 to 5 μM.

3.2.4 Analysis of spectroscopic data

3.2.4.1 Equilibrium data

Analysis of heat-induced and denaturant-induced unfolding curves followed the formalism describing a simple two-state transition between the folded and unfolded state, N and U, respectively. At each temperature or denaturant concentration, the observed signal S, representing either the molar ellipticity or the shift of the fluorescence emission maximum, is given by:

$$S = f_U S_U + (1 - f_U) S_N \tag{17}$$

where f_U is the fraction of unfolded molecule and S_U and S_N refer to the folded and unfolded state, respectively. S_U and S_N were assumed to be linear functions of temperature or denaturant concentration of the general form $S_i = S_{\sigma,0} + \alpha_\sigma T$ or $S_\sigma = S_{\sigma,0} + \alpha_\sigma$[denaturant], with σ indicating

the N or U state. For a given total concentration $[N]_0$, the equilibrium unfolding constant K_U is defined by:

$$K_U = \frac{f_U^n [N]_0^{n-1}}{1 - f_U} \prod_i^k m_i^{m_i} \tag{18}$$

where m_i is the stoichiometric coefficient of an unfolding subunit (if unfolding is coupled to subunit dissociation) and $n = \sum_i^k m_i$ is the order of the reaction. For the monomeric transition of the protein, $K_U = \frac{f_U}{1 - f_U}$. For the duplex DNA, $K_U = \frac{f_U^2 [N]_0}{1 - f_U}$. The unfolding enthalpy ΔH_m at the transition midpoint T_m was obtained from the temperature dependence of K_U according to:

$$K_U(T) = K_U(T_m) \exp\left\{ \frac{\Delta H_m}{R} \left(\frac{1}{T_m} - \frac{1}{T} \right) - \frac{\Delta C_p}{RT} \left[T - T_m - T \ln\left(\frac{T}{T_m} \right) \right] \right\} \tag{19}$$

R is the gas constant and ΔC_p is the heat capacity change. Eqs 17–19 can be combined to analyze thermal unfolding curves by non-linear curve optimization. Since T_m, ΔH_m and ΔC_p are strongly interdependent, the statistical significance of the values extracted from the analysis of a single trace is low. However, when ΔC_p is small compared to ΔH_m, the second term in the curly brackets of eqn 17 can be neglected.

In a different treatment, S_N and S_U were defined from the linear pre-transitional and post-transitional parts of the melting trace to obtain a plot of $f_U = \frac{S - S_N}{S_U - S_N}$ against T. From such plots, ΔH_m was calculated as [104]:

$$\Delta H_m = (2n + 2) R T_m^2 \left(\frac{\partial f_U}{\partial T} \right)_{T = T_m} \tag{20}$$

ΔC_p was obtained from plots of ΔH_m versus T (Kirchoff plots). To this end, thermal stability was varied by changing the solvent conditions. Knowing T_m, ΔH_m and ΔC_p, the unfolding free energy ΔG_U was calculated from the Gibbs-Helmholtz equation:

$$\Delta G_U(T) = \Delta H_m(T_m) \left(1 - \frac{T}{T_m} \right) + \Delta C_P \left[T - T_m - T \ln\left(\frac{T}{T_m} \right) \right] - RT \ln K_U \tag{21}$$

Data from isothermal denaturant titrations were analyzed following the linear extrapolation mode (LEM). The free energy of unfolding in the presence of denaturants, $\Delta G_U(D)$, is given by:

$$\Delta G_U(D) = -RT \ln K_U(D) = \Delta G_U(W) - m_D[D] \tag{22}$$

where m_D has units of kJ mol^{-1} M^{-1} and describes the linear dependence of $\Delta G_U(D)$ on [D]. $\Delta G(W)$ is the free energy of unfolding in the absence of denaturant. Since for a monomolecular unfolding transition K_U is 1 at $[D] = [D]_{1/2}$, we can write:

$$\Delta G_U(D) = m_D[D] - m_D[D]_{1/2} \tag{23}$$

The free energy of unfolding at zero denaturant concentration was calculated from $\Delta G_U(W) = m_D[D]_{1/2}$.

The association constant, K_A, from the fluorescence binding experiments was calculated from:

$$F = \frac{F_{\max}\left\{D_t + P_t + 1/K_A - \sqrt{(D_t + P_t + 1/K_A)^2 - 4D_tP_t}\right\}}{2D_t} \tag{24}$$

describing a 1:1 association reaction with D_t, P_t and F_{max} representing the total concentrations of DNA duplex and protein, respectively, and the maximal change of fluorescence intensity reached at saturation.

3.2.4.2 Kinetic data

Folding/unfolding experiments were analyzed by non-linear regression analysis according to

$$\ln k_{obs} = \ln(k_U e^{m_U D} + k_F e^{m_F D}) \tag{25}$$

where k_{obs} is the observed rate constant of the reaction, k_U and k_F are the unfolding and folding rate constants respectively, m_U and m_F are m-values of the unfolding and folding and D is the denaturant concentration. The ΔG was calculated from:

$$\Delta G = RT \ln\left(\frac{k_F}{k_U}\right) \tag{26}$$

assuming two-state mechanism of folding.

For the binding kinetics a fitting model for simultaneous determination of k_{on} and k_{off} was applied as reported elsewhere [105].

3.2.5 Differential scanning calorimetry

3.2.5.1 Experimental part

DSC experiments were performed on a VP-DSC calorimeter (MicroCal Inc.) equipped with twin coin-shaped cells of 0.52 ml volume. Details on the instrument's performance are given elsewhere [106]. The heating rate was 1 °C min^{-1}. Protein, DNA and complex samples were dialyzed for 18–24 hours against the same batch of buffer used to establish the baseline. Reversibility was checked by 2–3 cycles of heating and cooling. The raw experimental data were corrected for the instrumental buffer-buffer baseline and transformed to partial molar or partial specific heat capacity using partial specific volumes of 0.715 cm^3 g^{-1}, 0.561 cm^3 g^{-1} and 0.659 cm^3 g^{-1} for protein, DNA duplex, and protein-DNA complex, respectively (see below). The analysis of heat capacity traces of the protein-DNA complex followed the formalism detailed elsewhere [66, 107, 108]. Data handling and analysis were carried out using the program CpCalc 2.1 (Applied Thermodynamics), subroutines for Origin provided by MicroCal, and in-house written scripts for NLREG (Phillip H. Sherrod).

3.2.5.2 Analysis of DSC data

Deconvolution of the heat capacity traces followed the formalism described elsewhere [66, 107, 109]. Briefly, the excess enthalpy of a macromolecular system undergoing two-state unfolding, when defined as an "excess" over its native state, is:

$$\langle \Delta H \rangle = f_U \Delta H_{cal} \tag{27}$$

where ΔH_{cal} is the calorimetric molar enthalpy of denaturation of the cooperative unit. The excess heat capacity is obtained by differentiation of eqn 27:

$$\langle C_P \rangle = C_{P,N} + f_U \left(C_{P,U} - C_{P,N} \right) + \Delta H_{cal} \frac{\partial f_U}{\partial T} = C_{P,N} + f_U \Delta C_P + \Delta H_{cal} \frac{\partial f_U}{\partial T} \tag{28}$$

ΔC_p is the heat capacity change of unfolding; $C_{p,U}$ and $C_{p,N}$ are the heat capacities of the unfolded and folded state, respectively. $C_{p,N}$ is a linear function of temperature of the form a + b×T; $C_{p,U}$ can be approximated by a weak parabolic function, c + d×T + e×T^2. The progress of unfolding (fraction unfolded) is linked to the equilibrium constant through eqn 18. After explicit differentiation of the last term on the right-hand side of eqn 28, the excess heat capacity is given by :

$$\langle C_P \rangle = C_{P,N} + f_U \Delta C_P + \frac{\Delta H_{cal} \Delta H_{vH}}{RT^2} \times \frac{f_U \left(1 - f_U\right)}{n - f_U \left(n - 1\right)} \tag{29}$$

where ΔH_{vH} is the van't Hoff enthalpy of the process. If ΔH_{vH} is temperature-independent in the narrow temperature range where unfolding takes place, ΔH_{vH} equals $-R\partial \ln K_U / \partial(1/T)$ according to eqn 19. For cooperative two-state unfolding $\Delta H_{cal} = \Delta H_{vH}$. The combined equations 18, 19, 28 and 29 were fitted to the experimental heat capacity data to calculate the thermodynamic unfolding parameters characterizing a cooperative unfolding unit. For the general case of *i* overlapping, formally independent unfolding processes the apparent heat capacity is expressed as:

$$\langle C_P \rangle = C_{P,N} + \sum_i^k f_{U,i} \Delta C_{P,i} + \sum_i^k \Delta H_{cal,i} \frac{\partial f_{U,i}}{\partial T} \tag{30}$$

The van't Hoff enthalpy was calculated from DSC data according to:

$$\Delta H_{vH} = \left(\sqrt{n} + 1\right) T_{\max} \sqrt{R\left(C_{P,\max} - \frac{\Delta C_P \sqrt{n}}{\sqrt{n}+1}\right)} \tag{31}$$

where $C_{p,max}$ is the height of the heat absorption peak.

The heat capacity function was subjected to nonlinear regression analysis according to eqn 30, where *n* is the number of molecules dissociating in the melting zone. f_U can be numerically solved from:

$$K_C = \frac{f_U^n n^n C_{tot}^{n-1}}{1 - f_U} \tag{32}$$

where C_{tot} is the total concentration of complex and $K_C(T)$ is the calculated with reference to $T_{m0.5}$ ($f_U = 0.5$) by:

$$K_C(T) = 0.5^{n-1} n^n C_{tot}^{n-1} \exp\left[-\frac{\Delta H}{RT}\left(1 - \frac{T}{T_{m0.5}}\right) - \frac{\Delta C_P}{RT}\left(T - T_{m0.5} - T \ln \frac{T}{T_{m0.5}}\right)\right] \tag{33}$$

The absolute heat capacity C_p was derived from the apparent heat capacity, C_{app}, according to (*26*):

$$C_{app} = \left(C_p - \upsilon_M\right) m_M + const \tag{34}$$

where υ_M is the partial specific volume of the macromolecule and m_M is its mass in the calorimetric cell. The procedure requires the collection of thermograms at several different protein concentrations. The slope of plots of C_{app} vs. m_M numerically equals $C_p - \upsilon_M$. υ_M of the protein, 13 bp DNA duplex and the complex were computed according to Karshikoff &

Ladenstein [110]. This method reproduces with high precision υ_M of globular proteins. Indeed, the calculated υ_M is within 2% of υ_M estimated from the amino acid sequence according to Makhatadze et al. [111]. For the calculation, 20 NMR conformers of the complex and 25 NMR conformers of the protein were used and the DNA was modeled in an ideal B-form conformation. Values of υ_M used in eqn 34 were 0.561 g cm^{-3} for DNA, 0.715 ± 0.006 g cm^{-3} for protein, and 0.659 ± 0.003 g cm^{-3} for the complex.

3.2.6 Isothermal titration calorimetry

ITC was performed on a MCS ITC instrument (MicroCal Inc., Northampton, MA). The calorimeter was calibrated according to the manufacturer's description. Samples of protein and DNA were prepared in, and thoroughly dialyzed against, the same batch of buffer to minimize artifacts due to minor differences in buffer composition. Concentration was determined after dialysis. The sample cell (1.4 mL) was loaded with 15-70 μM DNA duplex. A titration experiment typically consisted of 20–25 injections of a 150–700 μM protein stock solution, each of 8 or 10 μL volume and 10 or 12 s duration, with a 5 min interval between additions. Stirring rate was 300 rpm. Non-specific heat effects were estimated from the magnitude of the peaks appearing after complete saturation. Raw data were integrated, corrected for nonspecific heats, normalized for concentration, and analyzed according to a 1:1 binding model assuming a single set of identical binding sites.

3.2.7 Binding simulations

To simulate the temperature dependence of the excess heat capacity function of an equimolar mixture of protein and DNA interacting to a 1:1 complex, the theoretical framework developed by Brandts and Lin was used [75]. In the following, SS_A and SS_B denote, respectively, the total molar concentrations of DNA single strands A and B. D_N and D_T are, respectively, free DNA duplex and total DNA in duplex equivalents. The total molar concentration of protein is P_T. The concentrations of free protein in the unbound native state and in the unbound unfolded state are designated P_N and P_U, respectively. The following expressions describe all the equilibria of the system. DNA melting with strand dissociation, $D_N \leftrightarrows SS_A + SS_B$, is described by the equilibrium constant $K_D = SS_A \times SS_B/D_N$, the enthalpy of DNA dissociation $\Delta H_D(T_D)$ at the reference temperature (melting temperature) T_D, and by the heat capacity increment of DNA unfolding $\Delta C_{p,D}$. The protein conformational equilibrium $P_N \leftrightarrows P_U$ is described by the equilibrium constant $K_P = P_U/P_N$, the protein unfolding enthalpy $\Delta H_P(T_P)$ at the reference

temperature T_P, and the protein unfolding heat capacity change $\Delta C_{p,P}$. The Gibbs free energies of DNA unfolding, ΔG_D, and protein unfolding, ΔG_P, are defined by the Gibbs-Helmholtz eqs 35 and 36:

$$\Delta G_D(T) = \Delta H_D(T_D)\left(1-\frac{T}{T_D}\right)+\Delta C_{P,D}\left[T-T_D-T\ln\left(\frac{T}{T_D}\right)\right]-RT\ln(0.85D_T) \tag{35}$$

$$\Delta G_P(T) = \Delta H_P(T_P)\left(1-\frac{T}{T_P}\right)+\Delta C_{P,P}\left[T-T_P-T\ln\left(\frac{T}{T_P}\right)\right] \tag{36}$$

Native protein and DNA duplex associate to form the 1:1 complex, PD, according to $D_N + P_N \leftrightarrows PD$. The association constant is $K_A = PD/(D_N \times P_N)$ and the change of the enthalpy (ΔH_A), entropy (ΔS_A), and free energy (ΔG_A) of association with temperature is given by:

$$\Delta G_A(T) = -RT\ln K_A = \Delta H_A(T_R)+\int_{T_R}^{T}\Delta C_{P,A}dT - T\left[\Delta S(T_R)+\int_{T_R}^{T}\Delta C_{P,A}d\ln T\right] \tag{37}$$

where T_R denotes an appropriate reference temperature.

Since SS_A and SS_B are always present in equimolar amount and since $SS_A = SS_B$ equals the concentration of unfolded DNA in duplex equivalents, mass conservation can be written as:

$$D_T = D_N + SS_A + PD = \frac{SS_A^2}{K_D} + SS_A + K_A P_N \frac{SS_A^2}{K_D} \tag{38}$$

$$P_T = P_N + P_U + PD = P_N + K_P P_N + K_A P_N \frac{SS_A^2}{K_D} \tag{39}$$

Substitution of $P_N = P_T/(1 + K_P + SS_A^2 \times (K_A/K_D))$ from eqn 39 into eqn 38 leads to an equation that can be numerically solved for SS_A. Once SS_A is known, the population of all other species at any temperature can be calculated from the above equations.

The excess enthalpy function of the system, using the complex as the reference state is:

$$\langle H\rangle = \Delta H_P\left(\frac{P_U}{P_T}\right)+\Delta H_D\left(\frac{D_U}{D_T}\right)-\left[\Delta H_A\left(\frac{P_N+P_U}{P_T}\right)\right] \tag{40}$$

and after differentiation, the excess heat capacity function is obtained.

3.2.8 Calculation of solvent accessible surface area

ASA was calculated using the program NACCESS with the default set of atomic radii and parameters [112]. The structures of free protein and free DNA in their binding competent conformation (protein* and DNA* in Scheme I) were generated by removing the protein or the

DNA from the coordinate file of the complex (1B69). The NMR structure of the free protein (2BB8) and of the 13 bp duplex DNA in a canonical B-DNA conformation were used to calculate the surface area changes associated with the structural rearrangement induced upon binding. Residues Ser 2, His 72, Asp 73 and Gly 74 missing in the NMR structure of the complex, and residues Ser 2 and Gly 74 missing in the NMR structure of the free protein were added manually and the resulting structures were minimized by standard protocols using CHARMM. ΔC_p ascribed to dehydration of the complex interface was calculated from:

$$\Delta C_P = a_{aliph}\Delta ASA_{aliph} + a_{ar}\Delta ASA_{ar} + a_{pol}\Delta ASA_{pol} \quad (41)$$

where the terms ΔASA (in units of $Å^2$) are the binding-induced changes in aliphatic, aromatic and polar surface, respectively, and the coefficients a_i are the elementary contributions (in units of kJ K^{-1} mol^{-1} $Å^{-2}$) to the heat capacity of hydration of the corresponding type of surface [49, 73, 113]. Likewise, the hydration enthalpy can be scaled from the amount and type of surface becoming inaccessible to bulk solvent when the complex is formed [50, 51]. Numerical values of the per-$Å^2$-contributions to ΔH were taken from Table XIII in Makhatadze & Privalov [49]. Other parameterization schemes that do not distinguish between aliphatic and aromatic surface contributions [51, 73] gave very similar estimates of ΔC_p. The surface calculations and the energy minimization procedures were performed by Dr. A. Abebe.

4 CONFORMATIONAL STABILITY OF THE PROTEIN

4.1 Conformational stability of protein deduced from denaturant unfolding

4.1.1 Equilibrium experiments

The protein contains a single tryptophan located at the DNA binding site. The fluorescence emission spectrum has a peak centered at 342 nm. Thermal and chemical denaturation shifts the emission maximum by 12-14 nm to the red and quenches the fluorescence (Figure 5).

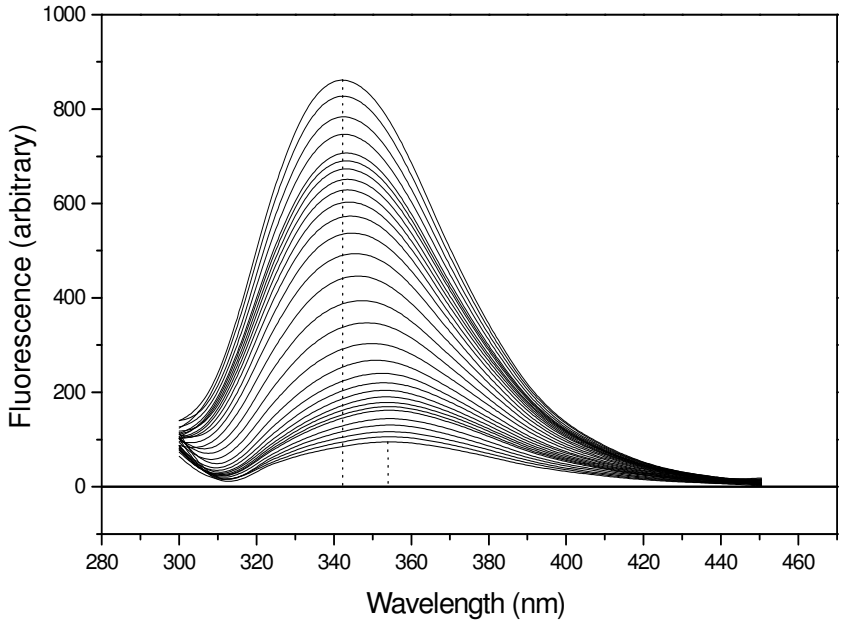


Figure 5. Fluorescence spectra of the protein in increasing concentrations of urea. The single tryptophane residue was exited at wavelength 295 nm and the spectra recorded from 300 nm to 450 nm. The topmost curve represents the fluorescence of the native protein and the lowermost the fluorescence of the fully unfolded protein at 7.5 M urea.

The shift of the emission maximum, λ_{max}, was used to follow denaturant induced unfolding (Figure 6). Unfolding by urea and GdmCl is well described by a two-state transition from the folded to the unfolded state (eqn 22, 23). The concentration of denaturant at the transition midpoint, $[D]_{1/2}$, progressively shifts to lower denaturant concentration as the temperature increases.

ΔG_U was calculated by the linear extrapolation method assuming a linear dependence of ΔG_U on the denaturant concentration. Values of ΔG_U from urea and GdmCl unfolding are very similar, implying no significant influence of (solvent screened) charge-charge interaction to the stability of the protein at pH 6 (Table 1 – page 101).

Values of m_D for GdmCl unfolding are the same within error between 12 and 27 °C whereas m_D from urea unfolding is slightly decreasing with increasing temperature (Figure 6). The m_D-value for denaturant-induced unfolding at constant temperature as well as the heat capacity change of unfolding are thought to correlate with the surface of protein exposed to solvent upon unfolding [49, 113, 114]. Even if this correlation has no rigorous thermodynamic foundation, it is a useful operational criterion to judge whether unfolding proceeds with a typical exposure of molecular surface. Assuming that the 74-residue DNA binding domain unfolds to a fully extended peptide chain, we calculate m_D-values of 4.12 kJ mol^{-1} M^{-1} for urea-induced unfolding and 8.72 kJ mol^{-1} M^{-1} from the amount of exposed surface [115]. The predictions

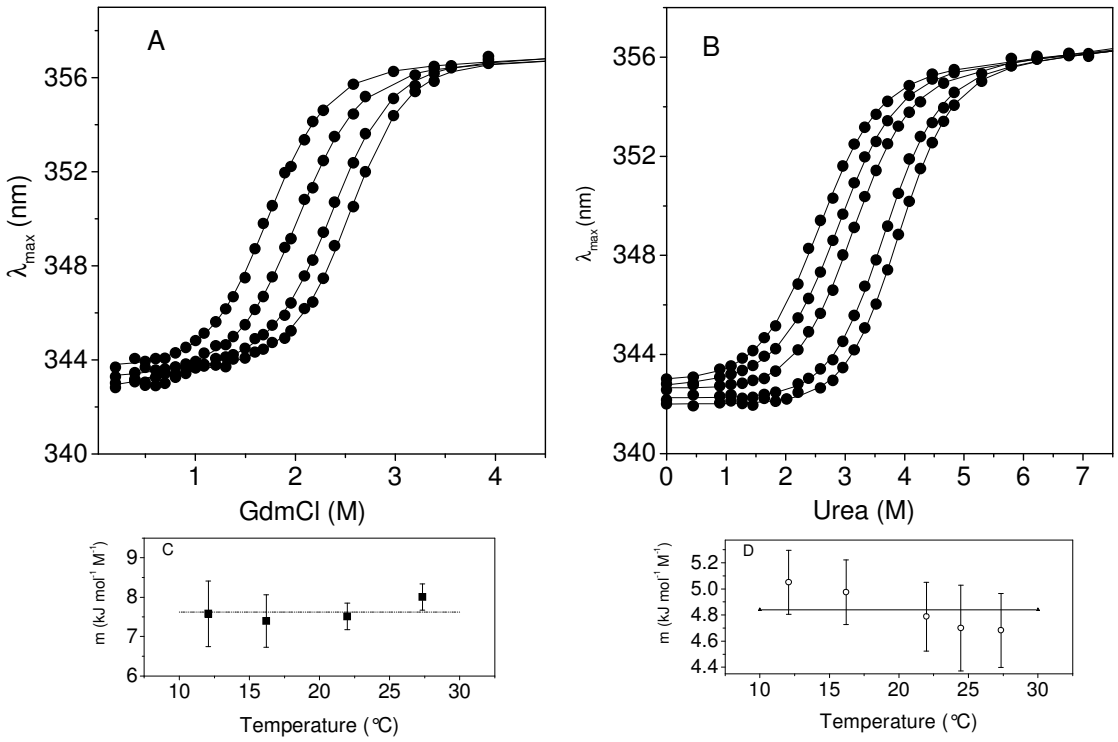


Figure 6. Protein unfolding induced by GdmCl (A) and urea (B) monitored by the shift of the tryptophan emission maximum. Symbols, experimental data; lines, best fits from non-linear regression analysis according to eqs 22, 23. Unfolding experiments were performed between 12 and 27°C (curves from left to right). (C) Slopes m_D from LEM analysis (eqs 22, 23) of GdmCl unfolding (filled symbols) and urea unfolding (open symbols). The dashed lines indicate the mean values.

correlate well with the mean m_D-values of 4.84 ± 0.33 and 7.62 ± 0.84 kJ mol^{-1} M^{-1} of Figures 6C and D.

4.1.2 Folding and unfolding kinetics

All the kinetic traces of the folding and unfolding of the protein at different denaturant concentrations were fitted satisfactorily by single-exponential function. In the refolding branch of the Chevron plot (Figure 7) the logarithm of the observed rate constant decreased from 4.33 to 1.53 upon increasing the urea concentration from 0.63M to 3.66M and a slope (m_F) of –0.97±0.02 M^{-1}. In the unfolding experiments the urea concentration was increased from 4.4M to 7.3M resulting in an increase of lnk_{obs} from 1.07 to 1.92 with a slope (m_U) of 0.67±0.04 M^{-1}. Fitting of the data according to eqn 25 yields values for the folding and unfolding rate constants of k_F = 138±6 s^{-1} and k_U = 0.05±0.01 s^{-1}. The concentration of urea at the midpoint, $[D]_{1/2}$, is 5M and the value for ΔG derived from the rate constants (eqn 26) at 7°C is 18.0±1 kJ mol^{-1}.The observed rate constants at different urea concentrations are reported in Table 2. The kinetic data from the unfolding and refolding of the protein was compared with those from the equilibrium unfolding experiments at different temperatures. The latter had to be extrapolated to 7°C for

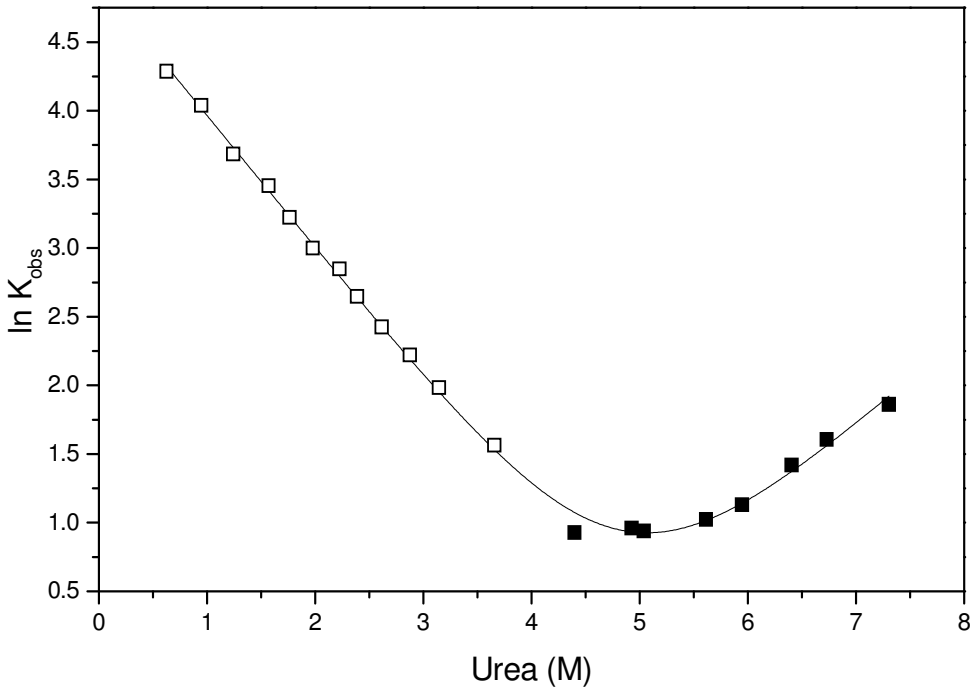


Figure 7. Observable rate of folding (open squares) and unfolding (solid squares) of INT-DBD. The line represents the best fit of the combined folding and unfolding data according to eqn.25.

proper comparison with the kinetic data. The value of ΔG = 18.0±1 kJ mol^{-1} derived from the rate constants (eqn 23) is slightly lower than ΔG from the equilibrium stability curve (Figure 15) calculated according to $\Delta G = m_D [D]_{1/2}$ - 20.9±1 kJ M^{-1} (from the equilibrium values of m_D and $[D]_{1/2}$ extrapolated to 7°C; see below) .

Other parameters from the kinetic experiments are also differing from the respective equilibrium values – the $[D]_{1/2}$ value from the Chevron plot is 5 M urea (Figure 7) while the linearly extrapolated $[D]_{1/2}$ at 7°C (from Table 1) is 4.03 M; the sum of the m – values for the unfolding and refolding experiments amounts to 3.81 kJ mol^{-1} M which is lower than the equilibrium m_D–value at 7°C of 5.19 kJ mol^{-1} M. These discrepancies suggest that the process of folding is not perfectly two state.

4.2 Thermal unfolding of protein followed by CD and fluorescence spectroscopy

4.2.1 Thermal unfolding followed by CD spectroscopy

Below 240 nm, the CD spectrum of the protein indicates contributions from α-helix, β-sheet, turns and loops scaled according to the relative content of these structure elements [116]. In the near-UV region from 260 to 300 nm there is a well-defined peak centered at 285 nm,

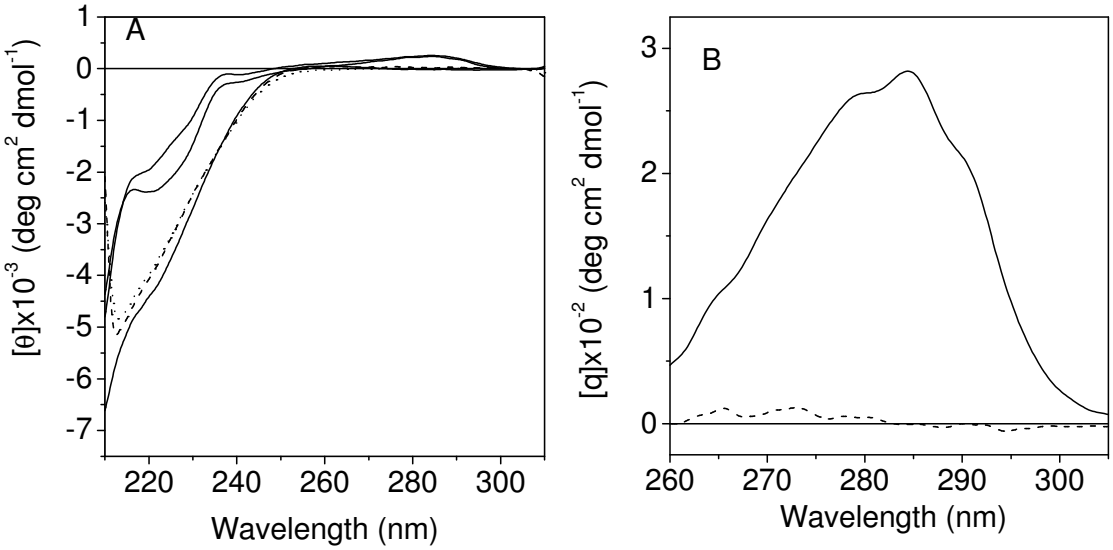


Figure 8. Circular dichroism spectra of the 74-residue N-terminal DNA-binding domain of bacterial integrase Tn916. (A) CD spectra measured with 65 μM protein in 50 mM Na-phosphate, 100 mM NaCl, pH 6.0 at (the solid lines from top to bottom at 220nm) 30 °C, 3 °C, 65 °C and in the presence of denaturing amounts of GdmCl (dash) or urea (dot). (B) Near-UV CD spectra measured in standard buffer at 30 °C (solid) and 65 °C (dash).

which mirrors the asymmetric environment of tyrosine and aromatic residues (Figure 8). Thermal unfolding as well as denaturant unfolding decreases the CD signal maximally at 219 nm (change of secondary structure) and 285 nm (change of tertiary structure in the environment of aromatic residues). The two wavelengths were used to follow thermal unfolding (Figure 9).

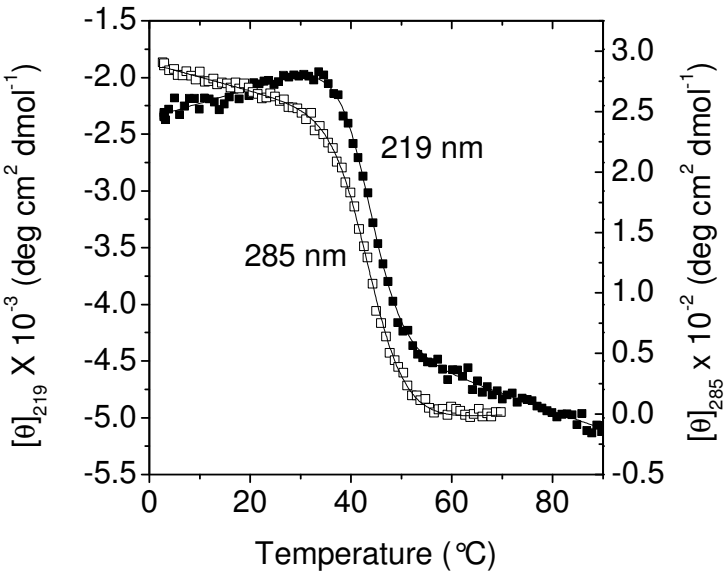


Figure 9. Thermal unfolding followed by circular dichroism. The data are presented as the change of the molar ellipticity per residue at 219 nm (filled symbols, left axis) and at 285 nm (open symbols, right axis). Lines are best fits for a two-state transition according to combined eqs 17-19.

The shape of the melting curves conforms to one major structural transition taking place in a relatively narrow temperature range. Below 30 °C the signal changes linearly at both wavelengths. This behavior is typical for many proteins and probably results from gradual structural changes in the native state. Above 60 °C the signal at 219 nm displays a temperature slope of about 14 deg cm^2 $dmol^{-1}$ K^{-1} while the signal at 285 nm remains constant. The spectral changes agree with a two-state unfolding mechanism (eqs 18 and 19). Values of T_m and ΔH_m deduced from the data at 219 nm and 285 nm are the same within error (Table 3). The difference between the two non-normalized melting curves of Figure 9 is due to the very different slopes before and after the main conformational transition. The difference disappears when the data are normalized to fraction of unfolded protein (Figure 15A).

4.2.2 Thermal unfolding followed by fluorescence spectroscopy

Figure 10 shows thermal melting curves represented as the shift of the fluorescence emission maximum after excitation at 295 nm. As for CD melting, the data agree with simple two-state transition from the folded to the unfolded state. Values of T_m and ΔH_m are identical within error to those deduced from the CD melting traces in Figure 9 (Table 3). Taken together,

the close correspondence of the midpoint temperatures T_m and the van't Hoff enthalpies ΔH_m deduced from the CD changes at 219 and 285 nm and from the shift of the fluorescence emission maximum very strongly indicates that the disruption of secondary and tertiary structure elements are tightly coupled processes conforming to a simple two-state transition reaction.

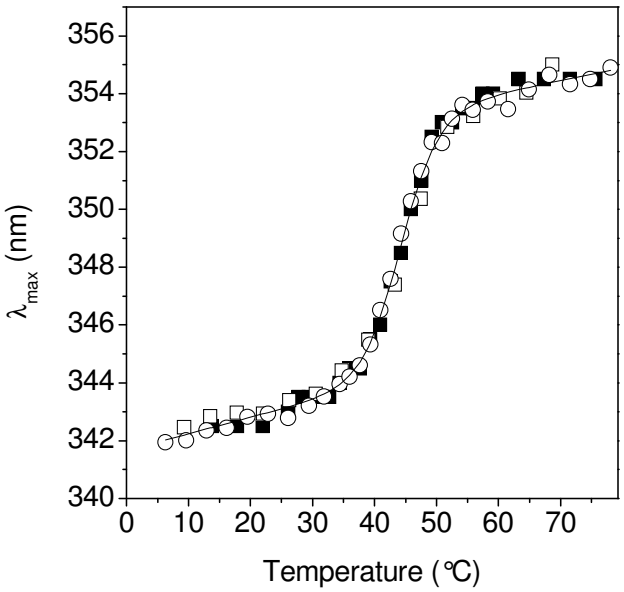


Figure 10. Thermal unfolding followed by fluorescence spectroscopy. The data are presented as the shift of the tryptophan fluorescence emission maximum. Different symbols refer to independent experiments. Line is best fit for a two-state transition according to the combined eqs 17-19.

4.2.3 Heat capacity change of protein unfolding

The heat capacity change of unfolding is defined as $\Delta C_p = \partial \Delta H_m / \partial T_m$. To estimate ΔC_P from the slope of a plot of ΔH_m against T_m (Kirchoff plot), variation of T_m and ΔH_m values in a broad temperature range had to be achieved. To this end, thermal unfolding experiments were repeated in the presence of small, non-denaturing amounts of GdmCl or urea. Figure 11 shows the Kirchoff plot obtained with the enlarged set of T_m and ΔH_m values (Table 4). The slope of the plot yields ΔC_p of 5.7 ± 0.5 kJ K^{-1} mol^{-1}.

The chosen procedure for obtaining ΔC_p may be flawed since the enthalpy and heat capacity of denaturant binding to proteins are negative and may affect the calculated thermodynamic parameters. The co-solutes may also influence the unfolding mechanism.
However, three observations indicate that the procedure is applicable in the present case. First, thermal melting curves recorded in the presence of non-denaturing amounts of urea or GdmCl are monophasic and can be fit by the same two-state transition model (Figure 12). Second, the progress of unfolding (f_U) is identical at any given pair of temperature and denaturant concentration when monitored by two different experimental probes, fluorescence and far-UV

CD spectroscopy (Figure 13). Third, ΔC_p calculated from the data of Figure 11 is in good agreement with ΔC_p deduced from the DSC experiments to be presented now.

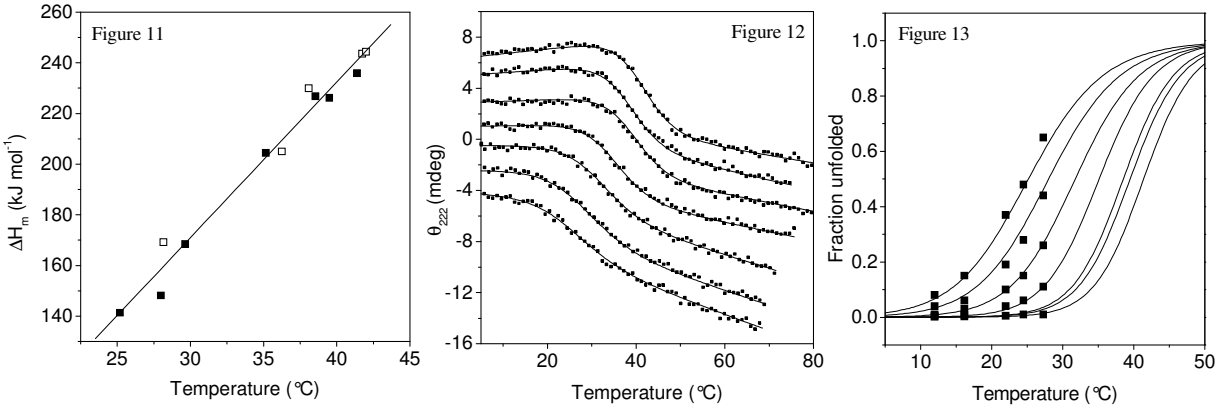


Figure 11. Determination of the heat capacity of unfolding, ΔC_p, from spectroscopic data. The slope is a linear best fit for ΔC_p of 5.7±0.5 kJ K^{-1} mol^{-1}. Filled symbols, data collected in the presence of urea; open symbols, data collected in the presence of GdmCl; see the text and Table 4 for details.

Figure 12. Thermal unfolding in the presence of non-denaturing concentrations of urea. Protein unfolding in the presence of 0 to 2.8 M urea (top to bottom) by monitoring the changes of the CD signal at 219 nm upon continuous heating. For clarity the traces are shifted on the vertical axis. Results from non-linear regression analysis according to the combined eqs 17-19 (see the main text) are presented by continuous lines. Values of T_m and ΔH_m are given in Table 4.

Figure 13. Fraction of unfolded protein as a function of temperature and urea concentration. The plot was constructed as follows: The apparent fraction of unfolded protein was calculated from the experiments shown in Figure 9, assuming that the linear signals observed before and after the main transition correspond to fully folded ($f_U = 0$) and fully unfolded ($f_U = 1$) protein, respectively. Lines from left to right correspond to the melting curves from 2.8 to 0 M urea. Likewise, the data from isothermal denaturant unfolding experiments shown in Figure 6B were normalized to calculate the fraction of unfolded protein at five fixed temperatures and at urea concentrations used in the experiments of Figure 11. The results are plotted as symbols.

4.3 Thermal unfolding of protein followed by DSC

4.3.1 Partial molar heat capacity.

Figure 14A shows the temperature-dependence of the partial molar heat capacity of the protein. The transition maximum and the shape of the calorimetric trace are independent of the protein concentration and the rate of heating, in agreement with a reversible monomolecular conformational transition. Hence, slow kinetic steps are not affecting the conformational transition. To eliminate possible complications caused by aggregation at high temperatures, DSC experiments were performed with low protein concentrations (typically 0.5 mg ml^{-1}) so that the absolute heat capacity C_p could not be measured precisely. Therefore, C_p was obtained from the

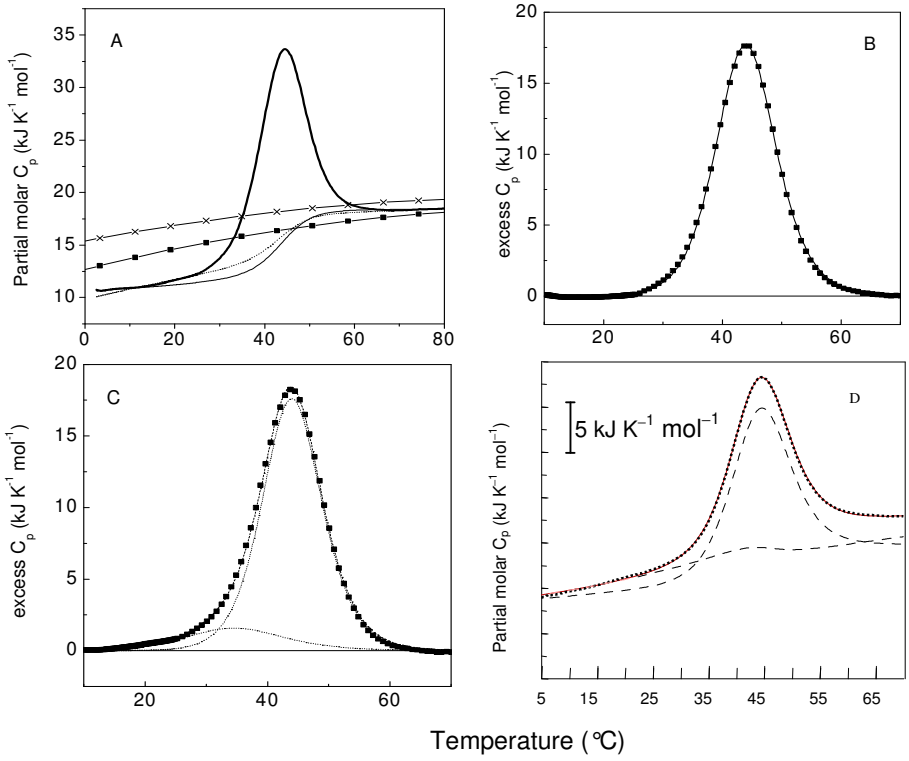


Figure 14. Thermal unfolding of the protein observed by DSC. (A) Temperature dependence of the partial molar heat capacity function (heavy line). The intrinsic heat capacity change is modeled in two ways: connection of the heat capacity of the folded state between 5 and 15 °C with the heat capacity of the unfolded state above 65 °C (thin solid line); connection of the heat capacity of the folded state between 15 and 25 °C with the heat capacity of the unfolded state above 65 °C (dotted line). The heat capacity of the denatured state calculated according to the HKF model (–■–) and the MP model (–×–) are also indicated. Experiments were performed with 65 μM protein in standard buffer of pH 6.0. (B) Excess heat capacity function calculated with the help of the dotted "baseline" of panel A. The experimental data (symbols) are best described by a single transition (i = 1 in eqn 30). (C) Excess heat capacity function calculated with the help of the solid "baseline" of panel A. The experimental data (symbols) are best described by two transitions (i = 2 in eqn 30). (D) Deconvolution of the apparent partial molar heat capacity. The two component reactions corresponding to a minor and a major transition are indicated by broken lines. The experimental function (solid line, corresponding to the trace of panel A) and the function synthesized using the two component reactions (dotted line) match perfectly. See the text for detailed description of the different analyses of the DSC traces.

change of the apparent heat capacity with protein concentration (eqn 34). The value of C_p at 25 °C is 12.25 ± 0.50 kJ K^{-1} mol^{-1}, corresponding to 1.43 ± 0.12 J K^{-1} per gram of protein. This is well within the range of specific heat capacities for other small globular proteins and equal to that of barnase [49, 67]. The initial slope of the pre-transitional portion of the specific heat capacity below 15 °C is small, 6–7×10^{-3} J K^{-2} g^{-1} but increases above ~15 °C to 12.5×10^{-3} J K^{-2} g^{-1}. This increase may point to a small conformational transition preceding the main unfolding reaction, as will be discussed below.

The heat capacity of the denatured protein is constant above 65 °C and has a value of 18.9 ± 0.7 kJ K^{-1} mol^{-1} (2.13 ± 0.08 J K^{-1} g^{-1}). The heat capacity of a denatured globular protein can be calculated by adding up the partial molar heat capacity of the constituent groups of the protein. Makhatadze and Privalov modeled the heat capacity of the amino acids of a fully solvated and unstructured protein by using data on small organic model compounds (model MP, [69]). Another recent procedure uses data on side chain analogues, poly-glycines and denatured proteins (model HKF, [68]). The expected heat capacity of the infolded protein calculated by models MP and HFK, respectively, have been added to Figure 14A. The experimental heat capacity trace of the protein is enveloped by C_p calculated by models MP and HKF above 65 °C.

4.3.2 Heat capacity change from DSC melting curve

The difference between the partial heat capacity of the unfolded and the folded state represents the direct calorimetric estimate of the unfolding heat capacity change, ΔC_p^{cal}. Taking the mean of the MP and the HFK model calculations, ΔC_p at 25 °C is 3.9 kJ K^{-1} mol^{-1} (53 J K^{-1} mol $residue^{-1}$) and 4.1 kJ K^{-1} mol^{-1} at 10 °C (56 J K^{-1} mol $residue^{-1}$). Considering the uncertainties of the model calculations, the two values of ΔC_p^{cal} are in agreement with ΔC_p of 5.7 ± 0.5 kJ K^{-1} mol^{-1} obtained from the Kirchoff plot of Figure 11.

The mean value of ΔC_p of 5.0 ± 0.8 kJ K^{-1} mol^{-1} (Table 3) agrees well with ΔC_p of 5.5 ± 0.8 kJ K^{-1} mol^{-1} calculated from the change of surface exposure according to different parameterization schemes [49, 73, 113]. One may therefore conclude that chemical and thermal denaturation of the DNA binding domain leads to a highly solvated state lacking significant intermolecular contacts. This is further supported by the fact that the heat capacity of the protein in the unfolded state matches the heat capacity predicted for a fully solvated polypeptide chain (Figure 15A).

By inference it follows that the folded DNA binding domain has a compactly folded structure, which is confirmed by packing density calculations performed on 25 NMR conformers. The calculated volume of voids and cavities of 2150 ± 60 and 13 ± 10 $Å^3$, respectively, are typical of a compact protein [110]. The specific void volume per atom is 1.80 ± 0.05 $Å^3$, similar to the median value for 35 proteins of molecular weight below 14 kDa [110]. The protein is also compact according to calorimetric criteria since the partial specific heat capacity at 25 °C resembles a typical globular domain [67]. However, starting from ~15 °C, C_p increases rapidly, indicating heat absorption by a temperature-dependent process. The changes in the environment of aromatic residues and the overall secondary structure also show

temperature dependence below 30 °C. Still, the heat capacity increase is not exceptionally steep and does not indicate substantial structural rearrangements and increasing surface exposure between 2 and 30 °C. In the NMR structure taken at 27 °C, close to the onset of the main heat absorption peak, the secondary structure elements are intact, the extended loops L1 and L2 are well defined and are anchored to the body of the protein, contributing to the hydrophobic core. DNA binding affinity is preserved up to 30 °C. Most likely, the initial slope of the heat capacity function reflects gradual accumulation of thermal motions.

MD simulations were performed for 2 ns in explicit water (Dr. Abebe) and results showed that on average, the protein is fluctuating more at higher temperature, particularly in its C-terminal half. Taken together, MD simulations make it likely that the increase in heat capacity above 15 °C reflects partial destabilization and increased thermal motions, perhaps preferentially in the C-terminal half.

4.3.3 Specific unfolding parameters

The thermodynamic parameters of unfolding are within the range of values for stable globular proteins [49]. At 43 °C the specific enthalpy of 3.5 ± 0.2 kJ (mol residue)$^{-1}$ and the specific entropy of 11.0 ± 0.7 J K^{-1} (mol residue)$^{-1}$ are among the highest reported for proteins (Table III of [49]). This applies also to lower temperatures since the specific heat capacity change is similar as for other proteins, 63 ± 11 J K^{-1} (mol residue)$^{-1}$. The unfolding free energy is rather low, about 14 kJ mol^{-1} at 25 °C (Figure 15B), and is comparable to ΔG_U of the SH3 spectrin domain (*29*). It appears that strong packing interactions are largely compensated by unfavorable entropic factors so that the stability of the DNA binding domain is low. Whether low stability is an intrinsic feature of the DNA binding domain of Tn916 integrase or only a feature of the isolated 74-residue DNA binding domain is not known. Interestingly, according to NMR data, the DNA-bound protein is more unstructured than the free protein, which is rather unusual. Low stability implies increased flexibility and facilitates local unfolding, hence, the low stability of the folded structure may have biological meaning [117, 118].

4.3.4 Minor conformational transition preceding the main unfolding reaction

Calculation of thermodynamic unfolding parameters from DSC data can be done in two ways. The “baseline” subtracted from the experimental data can be based on the heat capacity trace between 15 and 25 °C, the “baseline” marked by the dotted sigmoidal line of Figure 14A. Calculation of the excess heat capacity by subtracting this baseline from the main trace yields the trace shown in Figure 14B. Alternatively, we take as the heat capacity of the unfolded

protein the trace between 5 and 15 °C and subtract the corresponding "baseline" (thin solid line of Figure 14A) from the main trace to obtain the excess heat capacity curve shown in Figure 14C. Though the traces of Figure 14B and C appear similar, close inspection reveals distinct differences. Analysis of the trace of Figure 14B fits perfectly to a single transition (best fit for i = 1 in eqn 30) and ΔH_{cal} of 238 ± 12 kJ mol^{-1}. Analysis of the trace of Figure 14C fits best to two transitions (best fit for i = 2 in eqn 30) as indicated by the dotted curves with midpoint temperatures of 34.8 °C and 44.1 °C, respectively. The enthalpy change ΔH_m^{cal} defined by the area under the peak of Figure14C is 255 ±13 kJ mol^{-1}.

A second way to calculate the enthalpy of unfolding is by performing a van't Hoff analysis according to eqn 31. In the case of a single transition between the folded and unfolded state, the ratio $\Delta H_m^{vH}/\Delta H_m^{cal}$ has to be 1. Indeed, $\Delta H_m^{vH}/\Delta H_m^{cal}$ is 1.04 ± 0.07 for the excess heat capacity trace of Figure 14B, in accord with a single transition. However, in the case of Figure 14C, $\Delta H_m^{vH}/\Delta H_m^{cal}$ is 0.92 ± 0.07. Although this ratio is perhaps not significantly below unity, it indicates that the unfolding reaction between 15 and 65 °C shown in Figure 14C is not perfectly two-state and is best described by a minor and a major transition.

A more rigorous analysis according to eqn 30 makes no assumption about the "baseline" connecting the pre- and post-transitional heat capacity traces but uses a general value for the heat capacity function of a folded protein ($C_{p,N}$ of eqn 25). The mean temperature dependence of $C_{p,N}$ of stable small globular proteins is $6.7 \pm 1.0 \times 10^{-3}$ J K^{-2} g^{-1} (set of 12 proteins; Table III in [67]). Analyzing the trace of Figure 14A with this value of $C_{p,N}$ (recalculated for the molecular mass of INT-DBD), one obtains the minimal model shown in Figure 14D, which accurately describes the experimental heat capacity trace by two transitions. The main transition has a T_m of 43.7 ± 0.3 °C and a ΔH_m of 245 ± 19 kJ mol^{-1}. The minor transition has a T_m of 28.1 ± 5.0 °C and a ΔH_m of 29.7 ± 10.4 kJ mol^{-1}. ΔC_p of the main transition is 2.6 ± 0.5 kJ K^{-1} mol^{-1} and of the minor transition 1.9 ± 0.1 kJ K^{-1} mol^{-1}.

Values of T_m, ΔH_m and ΔC_p obtained from the different analyses of the DSC traces in Figure 14 as well as from thermal unfolding monitored by CD and fluorescence spectroscopy in Figures 9 and 10 are put together in Table 3. The thermodynamic parameters deduced from independent calorimetric and spectroscopic experiments agree well. Thermal denaturation of the DNA-binding domain of integrase Tn916 follows the two-state mechanism of unfolding closely but not perfectly. There is firm calorimetric evidence for a temperature-induced, partial conformational change starting about 25°C below the main cooperative unfolding reaction and

proceeding with a relatively low heat absorption, yet with a large heat capacity increase. The mean value of ΔC_p of unfolding is 5.0 ± 0.8 kJ K^{-1} mol^{-1}. This value was used to analyze the thermodynamics of protein-DNA association described in the next chapter.

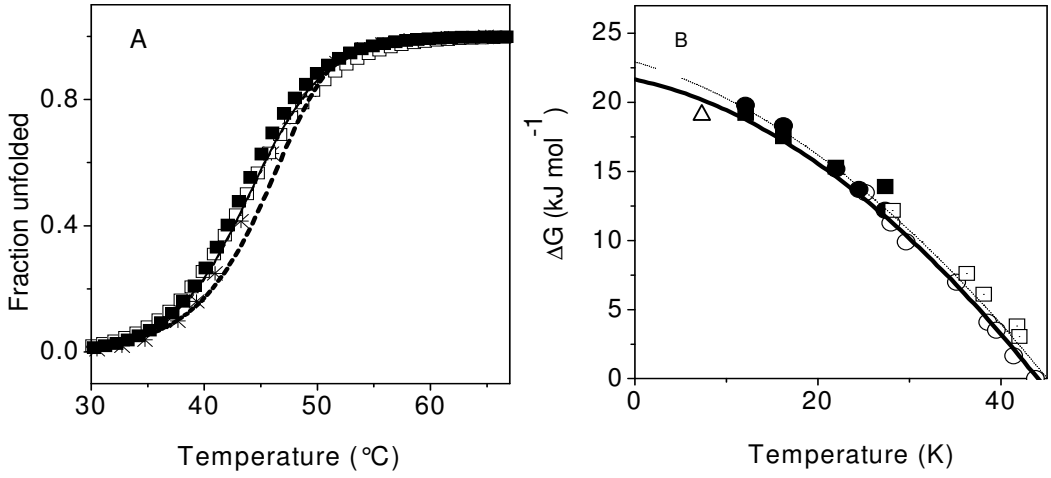


Figure 15. Thermal stability of the integrase DNA binding domain. (A) Fraction of unfolded protein as a function of temperature. Symbols refer to thermal unfolding followed by fluorescence (asterisks), CD signal at 219 nm (filled squares) and CD signal at 285 nm (open squares). Solid line, progress of unfolding calculated from DSC data in Figure 14C, including two transitions. Broken line, progress of unfolding calculated from DSC data in Figure14B, including only a single transition. (B) Thermal stability curves. Symbols: isothermal unfolding by urea (closed circles), isothermal unfolding by GdmCl (filled squares), thermal unfolding in the presence of non-denaturing amounts of urea (open circles, Table 4), thermal unfolding in the presence of non-denaturing amounts of GdmCl (open squares, Table 4). The lines illustrate best results from non-linear least-squares regression analysis according to the Gibbs-Helmholtz eqn 21. Heavy line, simulation based on the mean values measured by DSC (Table 3): T_m = 43.9 °C, ΔH_m = 250 kJ mol^{-1}, ΔC_p = 4.6 kJ K^{-1} mol^{-1}. Dotted line, combined urea and GdmCl unfolding data: T_m = 45 °C, ΔH_m = 267 kJ mol^{-1}, ΔC_p = 4.3 kJ K^{-1} mol^{-1}. The open triangle represents the stability derived from the kinetic data.

The DNA binding domain follows closely but not perfectly the two-state model of unfolding. Figure 15A shows that the heat-induced shift from the native to the unfolded state exhibits the same temperature dependence when monitored by the progress of heat uptake (DSC), the melting of secondary structure elements (CD signal at 219 nm), and the disruption of tertiary packing interactions (CD signal at 285 nm and tryptophan fluorescence emission). Good correspondence between DSC-derived and spectroscopy-derived parameters is achieved if a minor pre-transition with T_m around 30 °C is considered, but not if only the main transition is considered (compare solid with dashed line in Figure 15A). Combination of data from denaturant-induced unfolding and from unfolding by heat give a consistent picture of the energetics of the unfolding process. Figure 15B shows the protein stability curve constructed with the help of eqn 36.

5 CONFORMATIONAL STABILITY OF THE TARGET DNA

The temperature-induced unfolding and refolding of the 13 bp target DNA duplex was characterized by CD spectroscopy and DSC. The melting transition was highly reversible as inferred from the reproducibility of both the heat capacity trace and the ellipticity changes upon cooling and reheating of the sample. The CD spectrum of the DNA is typical for short synthetic DNA duplexes and displays reversible temperature-induced changes (Figure 16). Figure 16 shows thermal melting of the DNA monitored at 251 and 283 nm where the spectral changes are maximal. The data are well described by the van't Hoff formalism for bimolecular two-state unfolding (eqs 17-20). The same enthalpy of unfolding is calculated from the 251 nm and the 283 nm melting trace.

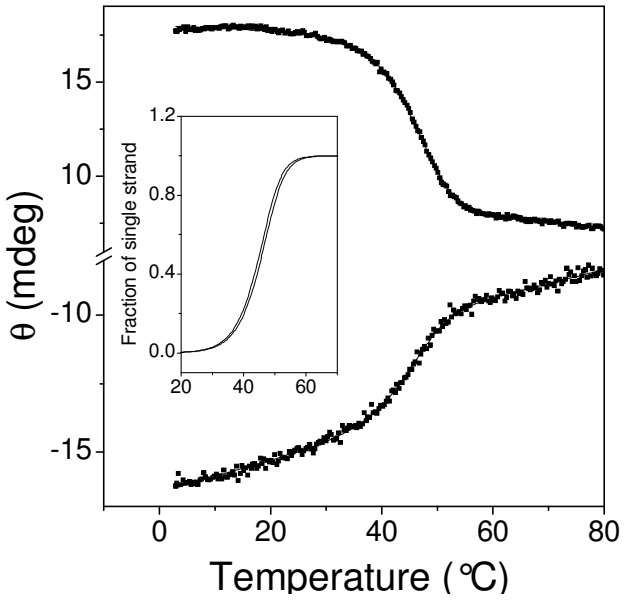


Figure 16. Thermal unfolding of the duplex DNA followed by CD. The change of ellipticity of a 65 μM DNA sample was monitored at 251 nm (lower trace) and 283 nm (upper trace). Continuous lines associated with the corresponding symbols are best non-linear regression fits according to equations 17-20. The inset shows the fraction of unfolded duplex calculated from the two experiments.

Figure 17A presents the partial molar heat capacity function of DNA melting obtained by DSC. The heat capacity peak develops from a rather low temperature and is asymmetric. The latter feature is typical for unfolding linked to dissociation. Temperature-induced strand dissociation is also evident from the increase of T_m with increasing DNA concentration. Linear extrapolation of the initial and final heat capacities into the transition zone (dotted lines in Figure 17A) shows a very small difference at T_m between the heat capacity of the duplex and the sum of the heat capacities of the unfolded complementary strands. The mean apparent transition

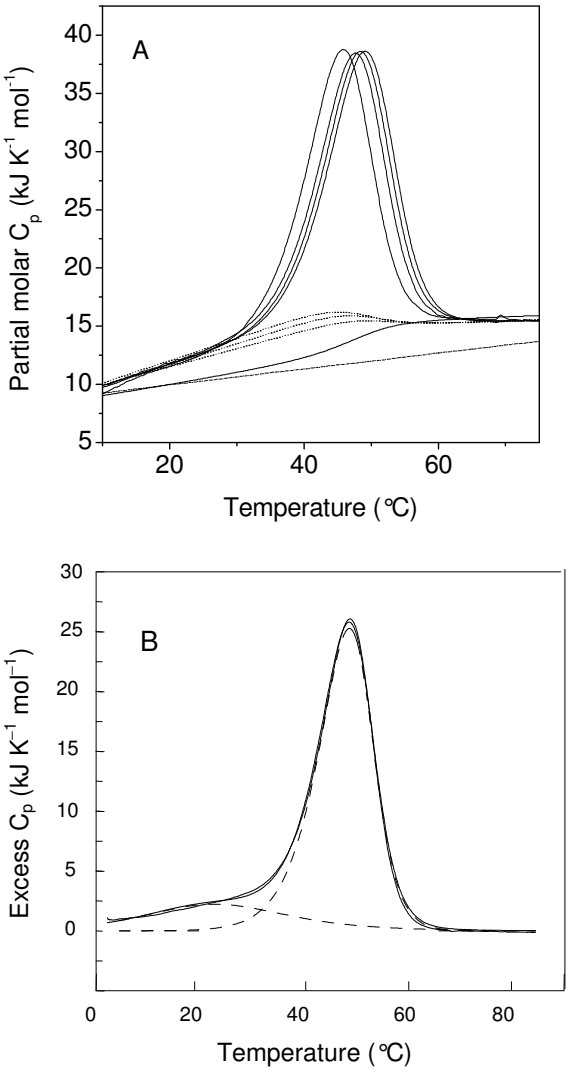


Figure 17. Thermal unfolding of the duplex DNA followed by DSC. (A) Melting of 33, 65, 90 and 119 μM duplex DNA (left to right) in standard buffer and at a scanning rate of 1 deg min^{-1}. Melting temperatures are 45.8, 47.6, 48.3, and 49.1 °C, respectively. Dotted lines connecting the pre-transitional and post-transitional traces model the apparent intrinsic heat capacity change. The calculated heat capacity of the folded duplex is indicated by the dash-dotted line (see text about how this line was constructed). (B) Excess heat capacity function calculated from the trace for 65 μM duplex DNA and the thin solid "baseline" of panel A. The low- and high-temperature components indicated by dashed lines correspond, respectively, to a broad pre-transitional accumulation of thermal energy and to cooperative melting accompanied by strand dissociation. Addition of the two components (dotted line) perfectly matches the experimental data (symbols).

enthalpy obtained by integration of the four heat capacity peaks delimited by the dotted "baselines" is 332 ± 4 kJ mol^{-1} (ΔH_{peak} in Table 5). A very similar value is obtained from plotting $1/T_m$ versus the concentration of DNA duplex (Table 5).

The heat capacity of the duplex increases steeply below the onset of the main transition. The average initial slope is 0.192 ± 0.017 kJ K^{-2} mol^{-1}, corresponding to 25×10^{-3} J K^{-2} per g of DNA. The latter value is four times higher than the mean value for stable globular proteins of 6×10^{-3} J K^{-2} g^{-1} [66]. Therefore, the experimentally observed pre-transitional heat capacity may not reflect the intrinsic heat capacity function of the fully folded DNA duplex. If so, the linear extrapolation of the pre-transitional heat capacity into the melting zone can not be used to estimate the heat capacity change characterizing unfolding and dissociation of the DNA duplex between low temperature and high temperature, say between 5 and 70 °C.

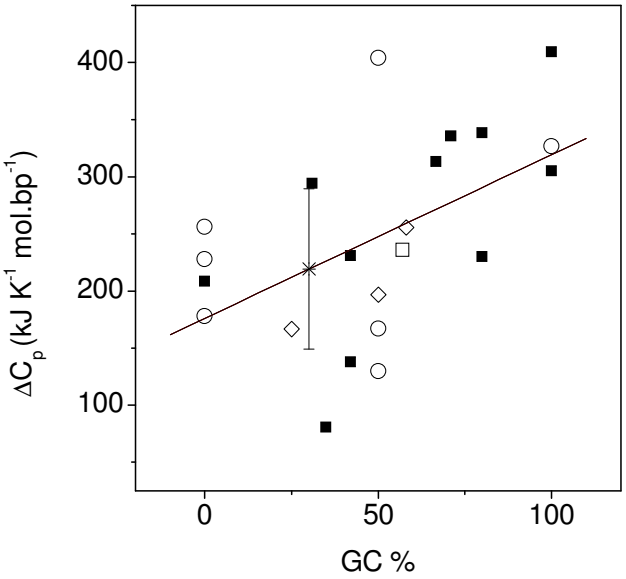


Figure 18. ΔC_p of 22 DNA duplexes vs their GC content. The 13 bp duplex used in this study has a GC content of 30 %. Its estimated ΔC_p with probable error is shown as an asterisk. ΔC_p values are from references [3] (open square), [5] (open diamond), [6] (open circle) and [7] (filled square).

The heat capacity change associated with DNA dissociation and unfolding has long been believed to be negligible because the heat capacities measured immediately before and after the main heat absorption peak of a thermogram are similar. On the other hand, Kirchoff analysis shows that the melting enthalpy is temperature-dependent [7]. Large ΔC_p values have been measured by isothermal titration calorimetry [3, 5, 119]. A more consistent picture of duplex melting has emerged only recently [6]. Isolated DNA strands form temperature-sensitive intramolecular structures, which are disrupted upon duplex formation. The heat capacity of short duplexes increases steeply with temperature and becomes approximately equal to the summed heat capacities of the separated strands. As suggested previously, this reflects temperature-dependent non-cooperative structural fluctuations of the duplex, e.g. end fraying and twisting as also seen in the present study (Figure 17).

An average ΔC_p value of 0.270 ± 0.089 kJ K^{-1} (mol bp)$^{-1}$ for DNA unfolding and dissociation was reported based on precise calorimetric analysis of polynucleotides [6]; a value of 0.196 kJ K^{-1} (mol bp)$^{-1}$ was obtained from Kirchoff's plots of ΔH_m vs. T_m from DNA melting data [6, 7]. We have combined the data published in references [3, 5, 120] and have corrected them for different GC content (Figure 18). From this large data set we calculate ΔC_p of 0.21 ± 0.07 kJ K^{-1} (mol bp)$^{-1}$, corresponding to ΔC_p of 2.7 ± 0.9 kJ K^{-1} mol^{-1} for unfolding and dissociation of the 13 bp target DNA duplex. Subtracting 2.7 kJ K^{-1} mol^{-1} from the heat capacity at 65 °C in Figure 17A one obtains the heat capacity of the folded duplex at that temperature. The straight line connecting this point with the heat capacity at the lowest temperature where the duplex is fully folded yields, as a first approximation, the intrinsic temperature dependence of the heat capacity of the folded state (dash-dotted line in Figure 17A). The slope of this function

is 69 ± 13 J K^{-2} mol^{-1} or (8.8 ± 1.7) $\times 10^{-3}$ J K^{-2} per g of DNA, in good agreement with (10 ± 2) $\times 10^{-3}$ J K^{-2} g^{-1} found for duplexes of 10, 12, and 16 base pairs [5].

With the help of the duplex heat capacity function so constructed, we can calculate the "baseline" linking the low and the high temperature limbs of the DSC trace (thin sigmoidal line in Figure 17A) and construct the excess heat capacity function shown in Figure 17B. Deconvolution analysis of the excess heat capacity function with the help of eqn 30 yields two different, partially overlapping phases: an initial gradual heat absorption, followed by a phase of intense heat absorption associated with the disruption of bulk packing interactions and bimolecular strand dissociation. The first phase is centered around 20°C and proceeds with absorption of 15-20% of the total heat. The enthalpies of the two processes are listed in Table 5.

The validity of the calculated heat capacity change is supported by three observations. First, the heat capacity function of the folded duplex derived with the help of ΔC_p of 2.7 kJ K^{-1} mol^{-1} (dash-dotted line of Figure 17A) is identical with the heat capacity function calculated for several other short duplexes [5]. Second, the enthalpy of the main dissociation and unfolding reaction (ΔH_{peak} of Table 5) changes by less than 10 % when ΔC_p is varied within its limit of uncertainty of ± 0.9 kJ K^{-1} mol^{-1}. Third, extrapolating the total melting enthalpy back to room temperature with the help of ΔC_p of 2.7 kJ K^{-1} mol^{-1} one obtains ΔH, ΔS and ΔG of duplex formation at 25 °C of 27 kJ $(mol\ bp)^{-1}$, 76 J K^{-1} $(mol\ bp)^{-1}$ and 4.1 kJ $(mol\ bp)^{-1}$, respectively. These values are typical for natural and synthetic DNA molecules. The dissociation constant of the 13 bp duplex DNA increases from 1×10^{-13} M at 5 °C to 1×10^{-8} M at 30 °C. At 30 °C and 30 μM concentration, the duplex is still more than 98 % populated.

Summarizing the analysis of the DNA duplex melting, we note that the enthalpy change of the main transition (ΔH_{peak} of Table 5) is well documented by independent calorimetric and spectroscopic experiments. As is typical of short DNA duplexes, there is a significant pre-melting phase (ΔH_{grad} of Table 5). Because of this gradual heat absorption below the main transition, ΔC_p could not be measured directly. Based on published results for many different DNA duplexes, we calculate ΔC_p of 2.7 ± 0.9 kJ K^{-1} mol^{-1} for the 13 bp duplex DNA. This value was used to analyze the thermodynamics of protein-DNA association described in the next chapter.

6 BINDING OF INTEGRASE TN916 TO ITS TARGET DNA

6.1 CD spectroscopy reveals conformational changes

The CD spectrum of the complex differs from the sum of the CD spectra of the protein and the DNA (Figure 19A). Subtraction of the CD spectrum at 85 °C from the spectrum at 3 °C yields the difference spectrum between the folded and the unfolded species (Figure 19B).

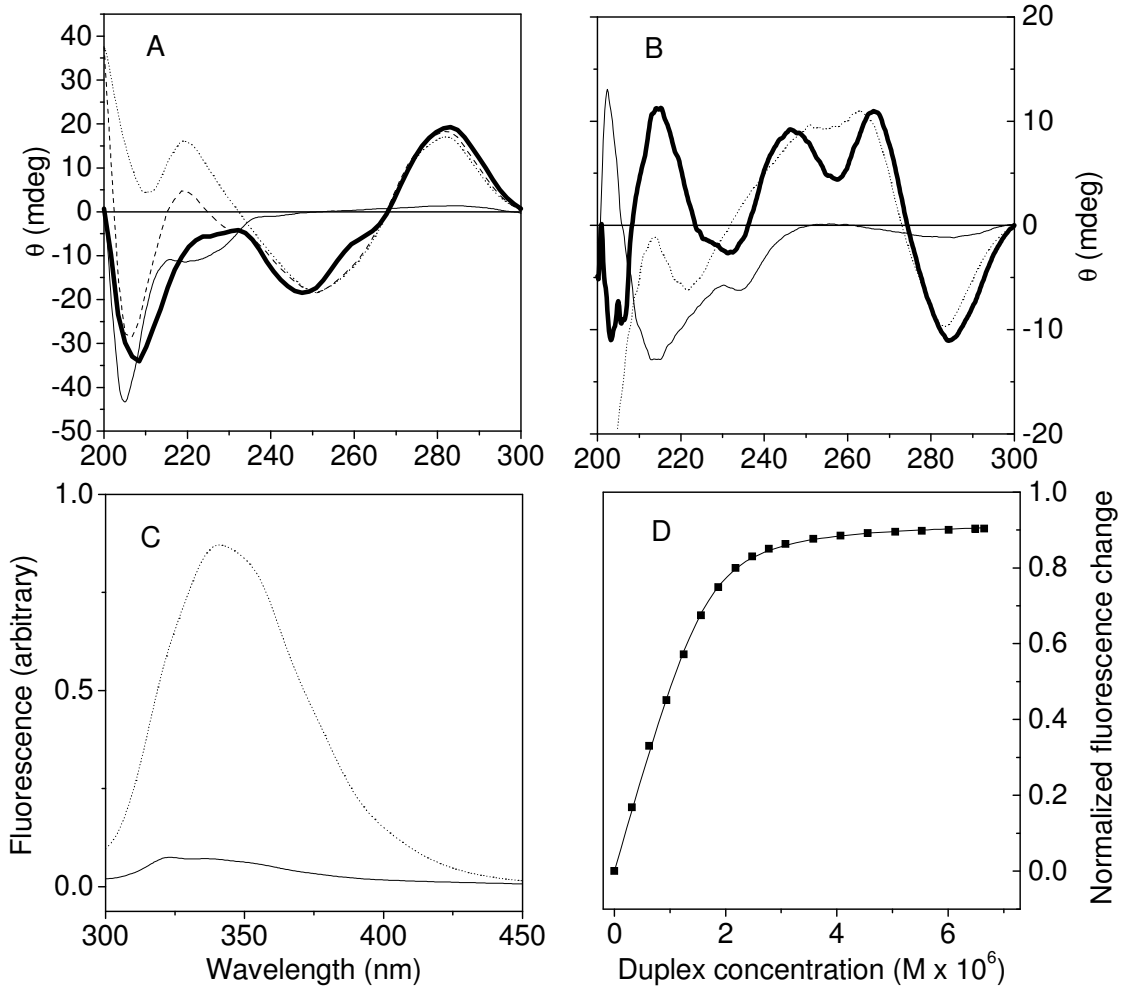


Figure 19. Spectroscopic characterization of protein, DNA and protein-DNA complex. (A) CD spectra of protein (thin line), DNA (dotted line), 1:1 protein-DNA complex (heavy line) and calculated sum of spectra of free protein and free DNA (dashed line) recorded at 3 °C in standard buffer with 60 µM protein, 60 µM duplex DNA, and 60 µM complex, respectively. (B) Difference spectra calculated by arithmetic subtraction of the spectra at 3 °C from the spectra at 85 °C (lines as in panel A). Spectral changes induced by heating are >95% reversible. (C) Fluorescence emission spectrum (excitation 295 nm) of free protein (dotted line) and of protein saturated with DNA (solid line). (D) Binding isotherm recorded at 20 °C. The data are plotted as the normalized change of the emission intensity. DNA from a 70 µM stock solution was added stepwise to 2 µM of protein. The continuous line is a best fit according to eqn 24 for K_A of 8.3×10^6 M^{-1}.

Thermal unfolding of the complex is silent at 232 nm for the DNA and at 260 nm for the protein. Ellipticities at these wavelength can be used to separately monitor the unfolding of the protein and the DNA, respectively, during thermal unfolding of the complex (Figure 20).

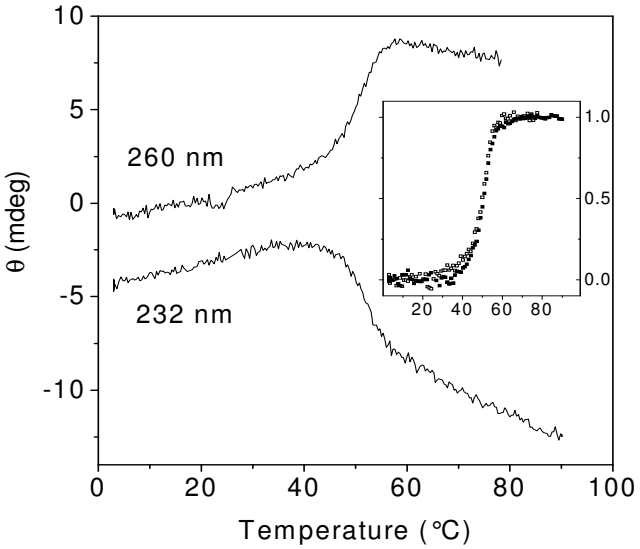


Figure 20. Thermal melting of the protein-DNA complex followed by CD. The change of ellipticity at 260 nm (melting of DNA) and 232 nm (melting of protein) was measured with 60 μM complex in standard buffer at a heating rate of 1 deg min^{-1}. Inset: Progress of reaction expressed as fraction of melted complex. Closed circles, melting of protein followed at 232 nm; open circles, melting of DNA followed at 260 nm.

6.2 K_D of the complex obtained by fluorescence titration of protein with DNA

The fluorescence emission of Trp 42, located at the DNA binding site [1], is quenched when this residue is buried at the complex interface (Figure 19C). From the change of fluorescence the association constant of the complex can be determined (Figure 19D). The solid line in Figure 19D is a best fit for a 1:1 binding reaction (eqn 24) with equilibrium constant K_A of $(8\pm3)\times10^6$ M^{-1} at 25 °C.

6.3 Kinetics of the complex formation measured by fluorescence stopped-flow

The rates of association and dissociation of the protein and DNA duplex were measured at 8 different concentrations of the complex components in the range of 160nM to 900nM. The average over all measurements gave: $k_{on} = (152\pm10)\times10^6 M^{-1}s^{-1}$ and $k_{off} = 11.2\pm3$ s^{-1}. The resulting K_D is 74±24 nM (Table 8), which is in the same order as the one obtained from equilibrium experiments (150nM – see below).

6.4 Thermodynamics of complex formation measured by titration calorimetry

The heat effect from the formation of the Int-DBD-DNA complex was measured with ITC. The mean of 20 experiments stoichiometry is 1.02±0.15, in agreement with a 1:1 complex. Within experimental error, K_A is temperature-independent between 4 and 30 °C. The average K_A

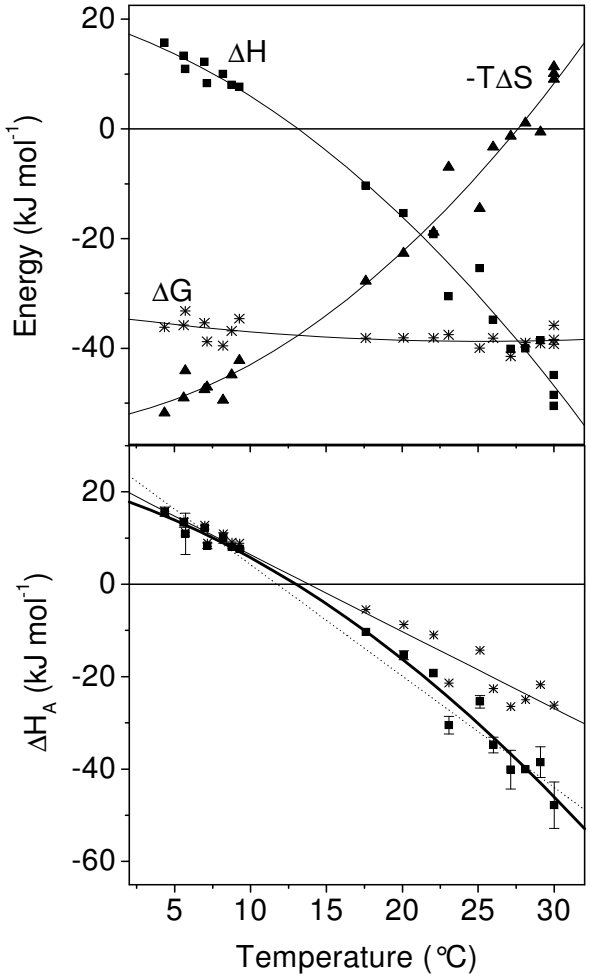


Figure 21. Energetics of protein-DNA association measured by ITC. (A) Changes of enthalpy, entropy and free energy in the range of 4–30 °C. (B) Determination of the heat capacity change $\Delta C_{p,A}$ from $\delta\Delta H_A/\delta T$ (Kirchoff plot). Filled squares, ΔH_A^{cal} (data from Table 6); asterisks, $\Delta H_{A,corr}$ (data from Table 6) for a hypothetical rigid body association reaction; see the text for detailed discussion. Dotted line: linear fit of $\delta\Delta H_A^{cal}/\delta T$ yields $\Delta C_{p,A}$ of –2.3±0.2 kJ K^{-1} mol^{-1}. Heavy line, second-order polynomial fit of $\delta\Delta H_A^{cal}/\delta T$ yields increasing $\Delta C_{p,A}$ from –1.4 kJ K^{-1} mol^{-1} at 4 °C to –2.9 kJ K^{-1} mol^{-1} at 30 °C. Thin solid line, linear fit of $\delta\Delta H_{A,corr}/\delta T$ yields $\Delta C_{p,A,corr}$ of –1.8±0.2 kJ K^{-1} mol^{-1} for a hypothetical rigid body association reaction.

is (7±5)×10^6 M^{-1}, in agreement with K_A from fluorescence titration (Figure 19D). The apparent enthalpy and entropy changes of association vary with temperature (Figure 21A). Complex formation is endothermic at low temperature and exothermic at high temperature, ΔH_A^{cal} changing sign at around 13 °C. The same ΔH_A^{cal} is measured in buffers of different heats of protonation indicating that there is no change in the protonation state of the protein or the DNA upon binding [50]. The entropy of complex formation is positive below about 29 °C and

negative at higher temperature (Figure 21A). The thermodynamic binding parameters from 20 ITC experiments performed in the range of 4 to 30 °C are summarized in Table 6.

The temperature variation of the binding enthalpy, $\delta\Delta H_A^{cal}/\delta T$, represents the heat capacity change of association, $\Delta C_{p,A}$. Linear regression of the data shown in Figure 21B yields $\Delta C_{p,A}$ of –2.3±0.2 kJ K^{-1} mol^{-1}. However, the data are not well described by a straight line. Fitting statistics improve significantly if a second-order polynomial is used (heavy line in Figure 21B). The curvature of the $\delta\Delta H_A^{cal}/\delta T$ function indicates a temperature variation from –1.4 kJ K^{-1} mol^{-1} at 5°C to –2.9 kJ K^{-1} mol^{-1} at 30 °C.

6.5 Thermal dissociation and unfolding of the protein-DNA complex

6.5.1 Thermal unfolding followed by CD

Unfolding was monitored at 232 nm and 260 nm, taking advantage of the observation that the two wavelengths reflect the thermal transition of protein alone and DNA alone (Figure 19B). The melting curves are shown in Figure 20. From the normalized curves (inset of Figure 20) it is seen that the transitions monitored at the two wavelengths are superimposable. This hints at a highly cooperative transition from folded complex to unfolded protein and DNA. The data of Figure 20 were subjected to van't Hoff analysis according to eqn 20. This analysis requires that the reaction order n of the thermal unfolding reaction is known. An order of 2 means that the complex unfolds into its folded components. An order of 3 means simultaneous dissociation and unfolding of the components. The enthalpy changes calculated from the CD melting curves, ΔH_m^{CD}, are 470 kJ mol^{-1} for n = 2 and 610 kJ mol^{-1} for n = 3. As will be shown below, the analysis of the DSC melting trace indicates that n has a value between 2 and 3.

6.5.2 Thermal unfolding followed by DSC.

The temperature-induced conformational transitions of the complex are highly reversible at pH 6 in 0.1 M NaCl. The partial molar heat capacity function is shown in Figure 22A together with the traces recorded for the isolated components. Melting of the complex produces a single sharp heat absorption peak whose temperature of maximum heat absorption (which can be regarded as the apparent transition temperature or melting temperature T_m) is higher than the melting temperatures of protein and DNA alone. The shape of the DSC trace of the complex is asymmetric, which indicates a cooperative unfolding process accompanied by subunit dissociation [108]. The DSC trace demonstrates that the components are thermally stabilized in

the complex (Figure 22B). The dissociation of the complex and the concurrent unfolding of the protein and the DNA are taking place within a narrow temperature interval.

The Heat Capacity Change at T_m Appears Negligible. The heat capacity of the complex increases linearly up to about 40 °C and remains constant above 60 °C, indicating that no further temperature-induced changes occur after the main transition (Figure 22A). Extrapolation of the pre-transitional and post-transitional heat capacities into the transition zone shows that the net heat capacity change of dissociation and unfolding of the complex, $\Delta C_{p,C}$, is almost negligible around T_m. In 7 experiments it scatters form –0.9 to +0.9 kJ K^{-1} mol^{-1} (dotted line in Figure 23). This is unexpected since $\Delta C_{p,C}$ is composed of the heat capacity change of protein unfolding ($\Delta C_{p,P}$), DNA unfolding ($\Delta C_{p,D}$), and of the dissociation reaction (equivalent to $-\Delta C_{p,A}$) according to $\Delta C_{p,C} = \Delta C_{p,P} + \Delta C_{p,D} - \Delta C_{p,A}$. $\Delta C_{p,P}$ of protein unfolding is 5 kJ K^{-1} mol^{-1} and $\Delta C_{p,D}$ of DNA unfolding is 1 to 2 kJ K^{-1} mol^{-1}. $\Delta C_{p,A}$ of complex association from ITC is –2 to –3 kJ K^{-1} mol^{-1} (Figure 21). Hence, the expected value of $\Delta C_{p,C}$ is 8 to 11 kJ K^{-1} mol^{-1}, in contradiction to the experiment.

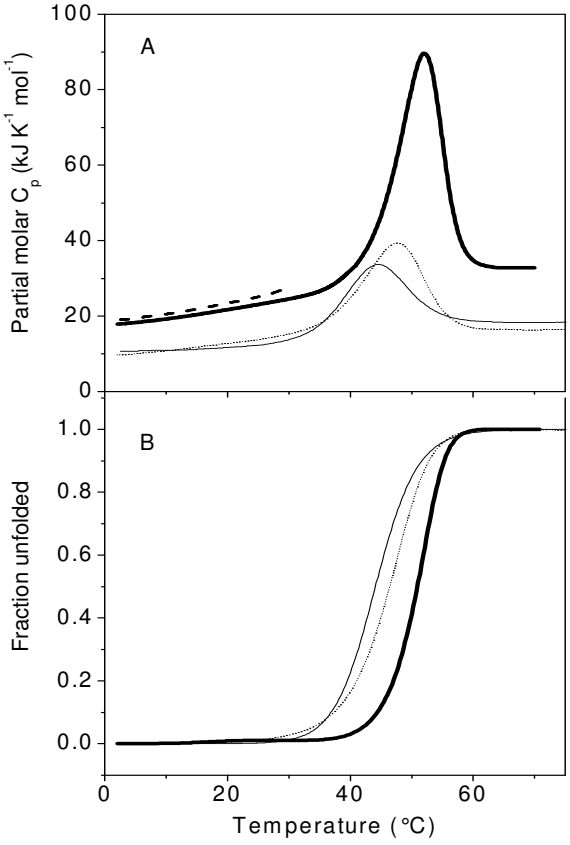


Figure 22. Heat capacity curves of the complex and the isolated components obtained by DSC. (A) Thermograms recorded in standard buffer of pH 6.0 at a heating rate of 1 deg min^{-1} with sample concentrations of 60 µM. Heavy line, complex; thin line, free protein; dotted line, DNA; dashed line, calculated sum of partial molar heat capacities of free protein and free DNA. (B) Progress of melting calculated from the changes of heat absorption. Line type as in panel A.

Why is the apparent $\Delta C_{p,C}$ deduced from the DSC melting curve so much smaller than expected? A possible explanation is that the complex dissociates into partly folded components so that the heat capacity is smaller than the sum of the heat capacities of the fully unfolded protein and the dissociated unstructured DNA strands. This interpretation can be excluded since the sum of the heat capacities of the protein and the DNA above 60 °C is within ±2 kJ K^{-1} mol^{-1}

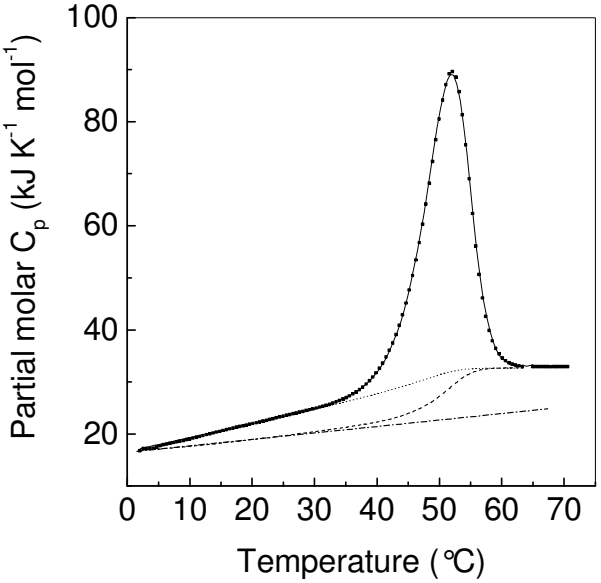


Figure 23. Partial molar heat capacity of the protein-DNA complex. Experimental conditions as in Figure 17. Symbols, experimental data. Continuous line, van't Hoff analysis according to eqs 31-33 with the following best fit parameters: n = 2.3, $T_{m0.5}$ = 51.1 °C, ΔH = 572 kJ mol^{-1} and ΔC_p = –0.15 kJ K^{-1} mol^{-1}. (Note that these parameters describe the experimental trace shown and differ from the mean of all experiments given in the text.) Dotted line, apparent intrinsic heat capacity function connecting the pre-transitional and post-transitional parts of the experimental trace and indicating negligible total heat capacity increment. Dashed line, intrinsic heat capacity of a hypothetical „non-fluctuating" complex undergoing thermal melting with a total heat capacity increment of 10 kJ K^{-1} mol^{-1}; see the text for detailed discussion. Dash-dotted line, heat capacity function of a hypothetical „non-fluctuating" complex.

of the heat capacity of the melted complex (Figure 22A). An alternative explanation may be hidden in the steep heat capacity increase of the native complex below 40 °C.

To further analyze this problem, we calculate the apparent calorimetric enthalpy of complex melting and dissociation, ΔH_C^{cal}. The value corresponds to the area above a line smoothly connecting the pre-transitional and post-transitional portions of the heat capacity trace (dotted line in Figure 23) and equals 540±20 kJ mol^{-1} (mean ± SD of 7 experiments). ΔH_C^{cal} is composed of the enthalpy changes of protein unfolding, DNA unfolding and complex dissociation (equivalent to $-\Delta H_A$): $\Delta H_C^{cal} = \Delta H_P + \Delta H_D - \Delta H_A$. Using the values determined for protein and DNA unfolding reported in [121] and ΔH_A from the ITC experiment extrapolated to the melting region, one obtains ΔH_C^{cal} of about 750 kJ mol^{-1}. This corresponds to the expected total heat change for the disruption of complex-stabilizing contacts and unfolding of the protein and the DNA. The value is significantly higher than ΔH_C^{cal} of 540 kJ mol^{-1} calculated from the

peak area of Figure 23. It follows that the DSC melting trace generates neither the expected enthalpy change nor the expected heat capacity change of complex dissociation and unfolding.

6.6 DSC study on the INT-DBD – DNA complex

6.6.1 Thermal fluctuations are gradually accumulating in the native complex before the start of the main melting transition

We propose that the discrepant values of measured and expected values of ΔH_C^{cal} and $\Delta C_{p,C}$ can be explained if one assumes that thermal fluctuations are gradually accumulating in the complex upon heating. This effect is illustrated in Figure 23. If, as discussed above, one assumes an expected total heat capacity change ($\Delta C_{p,C}$) of complex dissociation plus protein and DNA melting of 8 to 11 kJ K^{-1} mol^{-1}, the heat capacity of a hypothetical "non-fluctuating" native complex at 65 °C can be obtained by subtracting 8 to 11 kJ mol^{-1} from the high temperature end of the trace of Figure 23. This is shown by the dash-dotted line of Figure 23, which represents the heat capacity function of a hypothetical "non-fluctuating" protein-DNA complex lacking a steep increase of thermal fluctuation in the native state. The enthalpy change ΔH_C^{cal} of unfolding and dissociation of such a "non-fluctuating" complex corresponds to the area above the dashed line shown in Figure 23 and is 730 kJ mol^{-1}, in agreement with ΔH_C^{cal} of 750 kJ mol^{-1} calculated from $\Delta H_C^{cal} = \Delta H_P + \Delta H_D - \Delta H_A$. The slope of the heat capacity of the "non-fluctuating" complex is 125 J K^{-2} mol^{-1}, which is almost exactly the sum of the heat capacity temperature slopes of the DNA (69 J K^{-2} mol^{-1}) and the protein (55 J K^{-2} mol^{-1}).

Our analysis of the DSC melting trace indicates that the heat capacity of the complex increases rapidly with temperature. The enthalpic content of the system is significantly reduced at temperatures preceding the main transition. This indicates gradual destabilization of the complex. However, thermal fluctuations are linked to gain in entropy. Obviously, gradual enthalpic and entropic effects are effectively compensating each other, so that the Gibbs energy of association, ΔG_A, remains almost constant (Figure 21A).

6.6.2 Van't Hoff analysis of DSC trace confirms high cooperativity of complex dissociation and thermal unfolding

The melting profile of the complex shown in Figure 22A can be subjected to van't Hoff analysis. The reaction order n of eqs 31-33 is not defined. As discussed for CD melting above,

limiting values of n are 2 for the dissociation to the native components and 3 for the dissociation to unfolded protein and DNA strands. Simulation of the DSC trace according to eqs 32, 33 yields n = 2.3±0.1. The corresponding enthalpy change, ΔH_C^{vH}, is 560±20 kJ mol^{-1}, very similar to the measured ΔH_C^{cal} of 540±20 kJ mol^{-1}. Thus the ratio $\Delta H_C^{cal} / \Delta H_C^{vH}$ is 1, in support of a highly cooperative unfolding reaction. The enthalpies calculated for fixed n are 530 kJ mol^{-1} (n = 2) and 610 kJ mol^{-1} (n = 3). These limiting values are almost identical to those obtained from the CD melting curves (Figure 20): ΔH_m^{CD} = 470 kJ mol^{-1} (n = 2) and 610 kJ mol^{-1} (n = 3). Equally, fitting of the CD melting curves shown in Figure 20 with n fixed at 2.33 yields ΔH_m^{CD} of 525 kJ mol^{-1}, in good agreement with ΔH_C^{vH} of 560±20 kJ mol^{-1} and ΔH_C^{cal} of 540±20 kJ mol^{-1}.

The striking similarity of the enthalpy changes deduced from the CD melting curves, from van't Hoff analysis of the DSC trace and from direct calorimetric analysis strongly supports high cooperativity between complex dissociation and unfolding of the components. The reaction

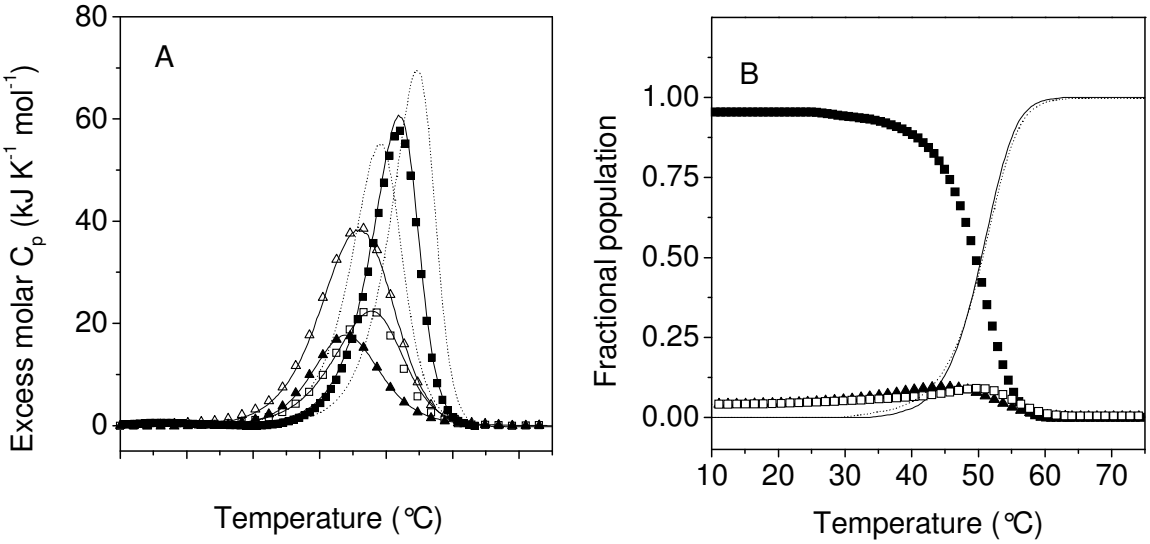


Figure 24. Simulations of protein-DNA association according to the procedure of Brandts and Lin [*70*]. (A) Symbols are the experimental excess molar heat capacity from DSC for the complex (filled squares), the free protein (open triangles), the free DNA (open squares), and the arithmetic sum of free protein and free DNA (open triangles). Solid lines are simulated excess heat capacity functions calculated with the help of the combined eqs 35–40 and the parameters $T_{m,\,protein}$ = 44.2 °C, $\Delta H_{m,protein}$ = 245 kJ mol^{-1}, $\Delta C_{p,protein}$ = 4.5 kJ K^{-1} mol^{-1}, $T_{m,DNA}$ = 46.1 °C, $\Delta H_{m,DNA}$ = 320 kJ mol^{-1}, $\Delta C_{p,DNA}$ = 1.5 kJ K^{-1} mol^{-1}, ΔH_A(25 °C) = –30 kJ K^{-1}, $\Delta C_{p,A}$ = –2.3 kJ K^{-1} mol^{-1}. The experimental melting temperature of the complex is reproduced best with K_A(25 °C) = 6.4×10^6 M^{-1}. Dotted lines are simulations performed with K_A one order of magnitude lower or higher than 6.4×10^6 M^{-1}. (B) Calculated temperature-induced changes of the fractional population of free native protein, free DNA duplex, and protein-DNA complex (symbols as in panel A). Solid line, free unfolded protein; dotted line, free single-stranded DNA.

order of 2.3, on the other hand, may indicate that the complex dissociates into components that may not yet be completely unfolded.

6.6.3 Simulation of the heat capacity trace by linked-equilibria analysis

The excess heat capacity function can be analyzed further by a model which takes into account the temperature-induced changes in the concentration of all the five molecular species populated in the temperature range of the DSC experiment: complex, native and unfolded protein, duplex DNA and unfolded DNA strands. This analysis combines the energetics of folding of the components described above with the energetics of complex formation described here. Although developed more than 10 years ago [75], the deconvolution technique was applied to protein-DNA interaction only once before [122]. Figure 24A presents the results of simulations according to the combined eqs 35–40. The experimental excess heat capacity profile of the 1:1 complex is very well simulated. The position of the heat absorption peak, reflecting the mutual stabilization of the individual components due to binding interactions, is reproduced with K_A(25°C) of 6.4×10^6 M^{-1}, which is the same within error as K_A of $(7\pm5)\times10^6$ M^{-1} from ITC and K_A of $(8\pm3)\times10^6$ M^{-1} from fluorescence quenching. ΔH_A^{cal}(25°C) determined by ITC is predicted within ±15 kJ mol^{-1}. Figure 24B shows the calculated population of each molecular species. In the transition region, melting of the complex releases small amounts of free folded protein and free DNA duplex in equilibrium with large amounts of unfolded protein and DNA. This is in accord with a reaction order between 2 and 3 obtained from van't Hoff analysis of the DSC trace. The success of the modeling indicates that the thermodynamic parameters obtained by ITC and DSC describe the thermodynamic behavior in a consistent way, bridging the entire temperature and concentration range covered by the ITC and DSC experiments.

6.7 The heat capacity change of protein-DNA association

6.7.1 Analysis of the partial molar heat capacities explains the $\Delta C_{p,A}$ temperature-dependence

A peculiar feature of the protein-DNA association reaction studied here is the nonlinear temperature-dependence of the association enthalpy seen in Figure 21B. We now wish to explain this behavior in terms of temperature-induced heat capacity changes occurring in the temperature interval where the native form of the complex and its components dominate. Such changes are clearly seen between 4 and 30 °C in Figure 22A. The calculated sum of the heat capacitiy slopes of protein and DNA (dashed line in Figure 22A) increases non-linearly from

0.24 kJ K^{-2} mol^{-1} near 4 °C to about 0.40 kJ K^{-2} mol^{-1} near 30 °C, the average over the ITC temperature range being 0.32 kJ K^{-2} mol^{-1}. At the same time, the heat capacity of the complex displays a nearly constant temperature slope of 0.24 kJ K^{-2} mol^{-1}. This behavior indicates the following: (i) In the temperature range before the main thermal transition, the free protein, the free DNA and the complex exhibit minor changes of conformation, of thermal fluctuations and of vibrational content, including changes in the water shell around the molecules. (ii) In the complex and its components the changes are not occurring in parallel. (iii) The enthalpy fluctuations of free protein and free DNA duplex are attenuated when association has taken place. These observations explain in a qualitative way $\delta\Delta H_A^{cal}/\delta T$ is curved (heavy line in Figure 21B).

If we were able to account in a quantitative way for the different thermal properties of the system in it's associated and dissociated states, we could predict the enthalpic behavior of a hypothetical rigid body association reaction between the protein and the DNA in their binding competent conformations (see Scheme I below). In other words, if we could subtract all the temperature-dependent (nonlinear) changes of C_p of the free components from the main trace of complex melting of Figure 22A, we would obtain the enthalpy change of the hypothetical rigid body association reaction. The necessary procedure has been developed by Privalov and colleagues [54]. The enthalpy of association at any temperature can be written as

$$\Delta H_A(T) = \Delta H_A(T_R) + \Delta C_{P,A}(T_R) \times (T - T_R) + \lambda \tag{42}$$

where λ is defined as

$$\int_{T_R}^{T} \left\{ \left[C_P(T) - C_P(T_R) \right]^c - \left[C_P(T) - C_P(T_R) \right]^i \right\} dT \tag{42a}$$

Superscript ^{i} denotes free protein and free DNA, respectively, and superscript ^{c} denotes the complex. $\Delta C_{p,A}(T_R)$ is defined as

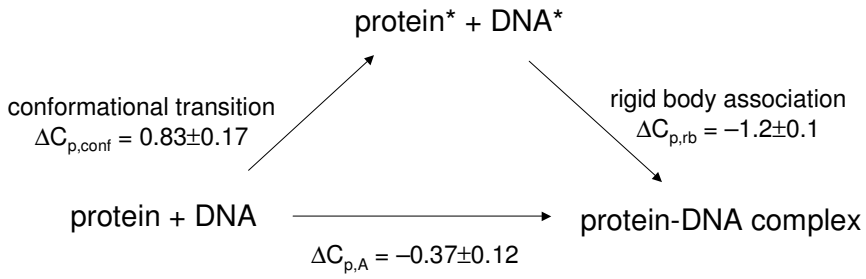
$$\Delta C_{p,A}(T_R) = C_p(T_R)^c - \sum_i C_p(T_R)^i \tag{42b}$$

Eqn 42 without the integral λ describes the ideal situation of a rigid body reaction for which $\Delta C_{p,A}$ is temperature-independent in the first approximation (in a sufficiently narrow interval of temperatures) and corresponds to the difference between C_p of the rigid complex and C_p of its rigid components. The integral λ represents the difference between the temperature dependence of C_p of the complex and the summed temperature slopes of C_p of free protein and DNA. Thus, we can define a "corrected" enthalpy change of association as

$$\Delta H_{A,corr} = \Delta H_A^{cal} - \lambda \tag{43}$$

Eqn 42 defines the association enthalpy change of the rigid body reaction. Values of $\Delta H_{A,corr}$ were calculated from the combined eqs 42 and 43. They are added as asterisks to Figure 21B and are shown in Table 6. Most satisfying, the change of $\Delta H_{A,corr}$ with temperature is linear, as expected for rigid body association. The corrected heat capacity change, $\Delta C_{p,A,corr} = \delta\Delta H_{A,corr}/\delta T$, is -1.8 ± 0.08 kJ K^{-1} mol^{-1} (slope of thin straight line in Figure 21B).

In conclusion, analysis of the heat capacities of free protein and DNA and their complex measured by DSC confirms the curvature of the plot of ΔH_A^{cal} versus T measured by ITC. The temperature dependence of $\Delta C_{p,A}$ is the consequence of non-parallel changes in the thermal fluctuations of the system in its dissociated and associated states. It is important to recognize that the "correction" of the ITC-measured enthalpy does not imply that $\Delta H_{A,corr}$ is the „true" binding enthalpy while the apparent experimental ΔH_A^{cal} is „wrong". Experimental ΔH_A^{cal} is the association enthalpy change of the molecules in their real conformational state, including thermal fluctuations, conformational changes and perhaps even partial unfolding. Our analysis merely highlights the thermodynamic complexity of the association reaction and provides a rationale for the nonlinear change of ΔH_A^{cal} with temperature. Furthermore, $\Delta C_{p,A,corr}$ of the hypothetical rigid body association reaction can now be judge against the structure of the complex, that means with regard to the molecular surface buried at the complex interface.



Scheme I

ΔC_p provides a semi-empirical link between structure and energetics and has attracted special attention as a key parameter in estimating the contribution of dehydration to binding [50, 51]. Experimental ΔC_p often correlates well with the amount and type of surface buried at the interface of a complex. However, in the case of protein-DNA association, significant

discrepancies between measured and calculated ΔC_p have been reported [45, 53-56]. The present system is a particularly telling example.

The conformational changes accompanying complex formation are reflected by the different solvent accessible surface area calculated for the complex and its free and bound components. The association reaction can be formally divided into two steps (Scheme I): (i) The transition of protein and DNA from the free to the binding competent conformations (marked by asterisks in Scheme I), and (ii) the association of the binding competent molecules to the complex.

The heat capacity changes can be estimated from the dehydration of nonpolar and polar surface area according to eqn 41. Estimates are based on empirical correlations established for many protein-protein and protein-DNA interactions [49-51, 73, 113]. Interestingly, the protein and the DNA unfold slightly to expose more binding surface in their binding competent conformation (+450 $Å^2$ of nonpolar surface and –80 $Å^2$ of polar surface). Therefore, $\Delta C_{p,conf}$ is positive with an estimated value of 0.83 kJ K^{-1} mol^{-1}. Association of the binding competent molecules buries 790 $Å^2$ of aliphatic, 540 $Å^2$ of aromatic and 1040 $Å^2$ of polar surface, leading to a negative value of $\Delta C_{p,rb}$ of –1.2 kJ K^{-1} mol^{-1}. The estimated overall heat capacity change of association is $\Delta C_{p,A} = \Delta C_{p,conf} + \Delta C_{p,rb}$ and is very much smaller than any of the measured values. Our estimates are based on elementary contributions per $Å^2$ of polar and nonpolar area (eqn 41) taken from [49]. Estimated values of ΔC_p differ by only 10-20% when other parameterizations are used [49-51, 73, 113] (not shown).

It should be emphasized that the calculated ΔC_p's are based on the simple assumption that the bulk contribution to ΔC_p originates from dehydration of interacting surface in the complex. Alteration of thermal motions and vibrational modi are not considered. Nevertheless, the structural estimates of ΔC_p provide us with a guideline for assessing the measured thermodynamic parameters.

6.7.2 Circumstantial evidence for residual water at the complex interface

Our analysis of the DSC data leading to $\Delta C_{p,A,corr}$ eliminates the contributions arising from conformational adaptation and thermal fluctuations. Therefore, $\Delta H_{A,corr}$ represents the enthalpy of inter-molecular contacts plus the changes of hydration upon association of rigid molecular surfaces. Hence, $\Delta C_{p,A,corr}$ should be equal to the structure-based estimate of $\Delta C_{p,rb}$. This is not the case: $\Delta C_{p,A,corr}$ is –1.8 kJ K^{-1} mol^{-1} and $\Delta C_{p,rb}$ is –1.2 kJ K^{-1} mol^{-1}. How can we explain the

"extra" –0.6 kJ K^{-1} mol^{-1}? We propose that they arise from incomplete dehydration of the complex interface.

It has long been recognized that water molecules play an important role in protein-DNA association [21]. Structural and computational analysis have identified water molecules bridging protein and DNA groups and often participating in hydrogen bonding networks [29]. Strong evidence for the proposed presence of water comes from packing density calculations. Using the NMR structures of complex, protein and DNA, we identify a total of 140±40 $Å^3$ of "empty space" distributed over 6 to 7 cavities at the complex interface. These cavities are large enough to contain together about 10 water molecules that are inaccessible to the bulk solvent. Since the "extra" heat capacity change is negative, it can be attributed to partial dehydration of polar surface (the heat capacity of polar hydration is negative). Thus, the –0.6 kJ K^{-1} mol^{-1} can be accounted for by assuming that the buried polar surface is only ~60% dehydrated. This estimation is based on an elementary contribution of hydration of –1.27 J K^{-1} mol^{-1} $Å^{-2}$ [49]. Dehydration is even less than 60% if one takes a value of –1.09 J K^{-1} mol^{-1} $Å^{-2}$ from the parameterization advanced by Freire and colleagues, [51].

6.7.3 Incomplete dehydration can be reconciled with the enthalpy of association

The association enthalpy measured by ITC is zero near 12-14 °C (Figure 21). The change from endothermic to exothermic heat effect close to room temperature indicates the importance of hydrophobic interactions since the enthalpy of apolar contacts is almost exactly cancelled by the corresponding dehydration enthalpy, hence the total enthalpic effect of hydrophobic interaction vanishes at 25 °C [49, 123]. The relatively small binding enthalpy in this temperature range results from weak intermolecular interactions, i.e. hydrogen bonds and other polar contacts, or from mutually compensating enthalpic effects. The enthalpy of an association process approximating rigid-body binding can be formally decomposed into four contributions:

$$\Delta H_A = \Delta H_{\text{int}}^{apol} + \Delta H_{\text{int}}^{pol} + \Delta H_{dehydr}^{apol} + \Delta H_{dehydr}^{pol} \qquad (44)$$

where the terms ΔH_{int} represent all the intermolecular contacts in vacuum and the terms ΔH_{dehydr} all the enthalpy effects of surface dehydration; superscripts ^{apol} and ^{pol} refer to contacts between nonpolar and polar surfaces, respectively. For energy parsing we apply the parameterization of Privalov and colleagues [49] because it clearly separates between vacuum energies and energy contributions from hydration based on calorimetric results. Using the total amount of buried surface to predict the total enthalpy effect of dehydration, we obtain about 1460 kJ mol^{-1} for $\Delta H_{dehydr}^{apol} + \Delta H_{dehydr}^{pol}$, of which about 90% arises from the very unfavorable enthalpy of burying

polar groups at the complex interface. The enthalpy of van der Waals contacts involving aliphatic and aromatic groups, ΔH_{dehydr}^{apol}, is only about –195 kJ mol^{-1}. Since the total association enthalpy of the rigid body association reaction at 25 °C, ΔH_A^{corr}, is –19 kJ mol^{-1} (Figure 21B), one obtains ΔH_{dehydr}^{pol} of –1270 kJ mol^{-1} from insertion in eqn 44. It is reasonable to assume that the bulk of ΔH_{int}^{pol} originates from hydrogen bonds at the complex interface. Again with the parameterization from [49], we take a value of –45 to –60 kJ mol^{-1} for the enthalpic content of a single hydrogen bond and predict about 21-28 hydrogen bonds at the complex interface. There are only 13 residues capable of forming H-bonds between the protein and the DNA. Hence the predicted number of H-bonds is too large. However, if only 60% of the buried polar surface were dehydrated upon binding, ΔH_{dehydr}^{pol} would be only about –760 kJ mol^{-1} and would agree with 13 to 17 H-bonds. This estimate is corroborated by NMR structural data [1] and MD simulations [15]. Thirteen H-bonds can be observed in at least 2-3 conformers of the NMR ensemble. Up to 16 hydrogen bonds can be identified in some conformers after optimization of the hydrogen positions. Ten residues are hydrogen-bonded to DNA bases and backbone groups throughout a 2 ns MD simulation. Furthermore, if water molecules are present at the interface, they also may serve as donors or acceptors of additional H-bonds. Indeed, we observe 7-8 water molecules inaccessible to bulk water in the simulations. Altogether, the enthalpy parsing analysis reconciles structural features of the complex with the observed heat capacity changes if one assumes that only 60% of the complex interface is dehydrated.

In concluding this discussion of enthalpy parsing, we note that polar interactions contribute only little to the enthalpy of binding around room temperature. Incidentally, the calculated enthalpy of hydrophobic bonding ($\Delta H_{int}^{apol} + \Delta H_{dehydr}^{apol}$), which includes van der Waals contacts between aliphatic and aromatic groups plus the corresponding dehydration of these groups, is –18 kJ mol^{-1} at 25 °C, which matches perfectly with the association enthalpy of the rigid body reaction at that temperature (thin line in Figure 21B). Thus, the analysis according to eqn 44 seems reasonable.

7 SALT AND OSMOTIC EFFECT ON THE INT-DBD TO DNA BINDING

7.1 Salt dependence of the stability of the Int–DBD and the DNA duplex.

To check for salt-dependent changes of protein and DNA stability, CD melting experiments were performed in 0.32 M [Na^+] (the highest salt concentration used in the ITC experiments), and the results were compared with stabilities at lower salt concentration [124]. The protein melts at 47 °C, ~3 °C higher than in low-salt buffer. The unfolding enthalpies in low-salt and high-salt buffers are identical within error. Assuming that the salt dependence of the unfolding heat capacity change is negligible in the narrow range of 0.15-0.32 M [Na^+], it follows that Int-DBD is marginally stabilized by high salt. The melting temperature of the DNA duplex in 0.32 M [Na^+] is higher by 5.5 °C as compared to melting in law salt. Stabilization is entropic in nature as expected [125].

7.2 Thermodynamics of complex formation in buffers with different salt concentrations.

7.2.1 Experimental results

Isothermal calorimetric titrations of Int-DBD into the DNA duplex were performed at 25 °C at varying Na^+ concentrations. The upper limit of [Na^+] was set at 0.32 M since the protein tended to aggregate at [Na^+] > 0.32 M at the high concentrations needed for ITC experiments in the weak-binding regime. The stoichiometry was 1 Int-DBD bound per 13 bp duplex at all salt concentrations tested. Representative experiments are shown in Figure 25 and thermodynamic parameters of binding are summarized in Table 9. Binding strongly depends on salt, the association constant, K_A, decreasing 20 times from 0.12 M to 0.32 M [Na^+] (Figure 26). This result is in line with the general behavior of the majority of sequence-specific protein-DNA complexes studied to date. In a most general form the binding equilibrium can be written as [126]:

$$P + D \leftrightarrows PD + cX^- + aM^+ + wH_2O \qquad (45)$$

where P, D and PD are the molar equilibrium concentrations of protein, DNA and complex; X^-, M^+ and H_2O are the molar concentrations of anions, cations and water, respectively; *c*, *a* and *w* denote stoichiometric coefficients. A quantitative relation has been proposed between the

decrease in the equilibrium constant, the increase of the salt concentration, and the number of ionic contacts at the complex interface [126]:

$$SK_{obs} = \left(\frac{\partial ln\, K_{obs}}{\partial ln[MX]}\right)_{T,p} \approx -\left[a + c - w\frac{2[MX]}{[H_2O]}\right] \quad (46)$$

The effect of anions (*c*) expelled from (mainly) the protein surface could not be generalized and is believed to be negligible. The last term in the above equation accounts for the increased osmotic pressure at high salt concentration, leading to sequestering of water molecules (if any

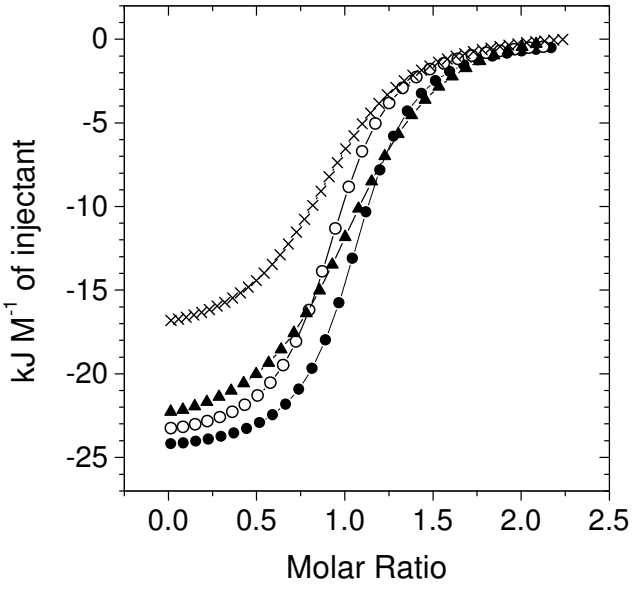


Figure 25. Binding isotherms measured by ITC for Int-DBD binding to the 13 bp DNA duplex. Titrations were performed at 25 °C in phosphate buffer, pH 6.0 supplemented with 0.12 M (filled citcles), 0.14 M (open circles), 0.19 M (triangles) and 0.32 M (crosses) NaCl. To facilitate comparisons the experimental data points are omitted and only the best non-linear least-squares fits to a 1:1 binding model are shown.

are present) from cavities at the complex interface. At low salt concentration where the osmotic effect of salt is negligible, and in the absence of an anion effect, the number of cations (*Z*) released from the DNA is obtained from the simplified form of eqn 46:

$$SK_{obs} = \left(\frac{\partial ln\, K_{obs}}{\partial ln[MX]}\right)_{T,p} = -Z\psi \quad (47)$$

The coefficient ψ equals 0.88 for helical B-DNA since it has been predicted that at low salt concentrations the ion gradient around the DNA neutralizes around 88% of its cumulative negative charge. However, newer models suggest that ψ = 0.64 is more appropriate in the case of short DNA duplexes [127, 128]. Using ψ = 0.64 and the data from Figure 26, the number of Na^+ cations released from the DNA upon complex formation is Z = 4.8±0.2.

The free energy, enthalpy and entropy are also linear functions of $[Na^+]$ with slopes $\frac{\partial \Delta G}{\partial \log\left[Na^+\right]} = 17.5 \pm 0.4$, $\frac{\partial \Delta H}{\partial \log\left[Na^+\right]} = 11.9 \pm 0.8$ and $\frac{\partial T\Delta S}{\partial \log\left[Na^+\right]} = -5.6 \pm 0.9$, respectively (units of kJ mol^{-1} $(logNa^+)^{-1}$, Figure 27). Obviously, the enthalpic contribution to the salt effect in the case of Int-DBD-DNA binding is significant. In fact, it dominates the observed change in free energy.

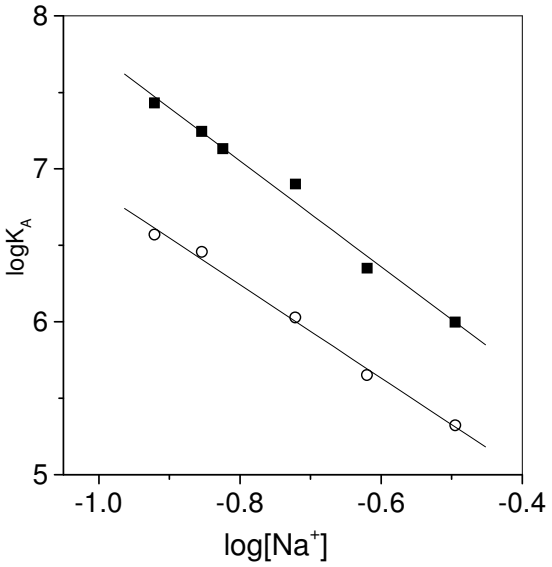


Figure 26. Salt dependence of the equilibrium constant of Int-DBD binding to DNA. Open circles are data obtained by ITC at 25 °C. Solid squares represent $logK_A = log(k_{off}/k_{on})$ calculated from the ratio of the rate constants for association (k_{on}) and dissociation (k_{off}). Kinetic data were collected by fluorescence stopped-flow experiments at 7°C. Continuous lines are linear fits according to eq 47. The slopes SK are −3.07±0.2 and −3.46±0.5 from ITC data and kinetic data, respectively.

The temperature dependence of the enthalpy change at the highest salt concentration of 0.32 M was measured at 4.5 °C and 25°C. The calculated heat capacity change, ΔC_p, is 1.8 ± 0.2 kJ K^{-1} mol^{-1}. This value is identical within error with ΔC_p we have previously obtained at low salt conditions [124]. Therefore, the increase of the ionic strength does not have a detectable effect on the value of the heat capacity change.

Titration experiments were conducted at the highest and lowest salt concentrations in buffers containing Br^- instead of Cl^-. Binding in the presence of Br^- is slightly weaker: $\Delta\Delta G = \Delta G^{Br} - \Delta G^{Cl}$ = 1.6 kJ mol^{-1} on average. The salt dependencies of K_A in NaBr and in NaCl are the same within error. Also the slopes of plots of ΔH versus $log[Na^+]$ are quite close (Figure 27). However, the enthalpy-entropy partitioning of ΔG is different as seen in Figure 27. At the same salt concentration substitution of Cl^- by Br^- results in an enthalpic penalty of ~2.7 kJ mol^{-1} and an entropic gain of ~1.1 kJ mol^{-1}.

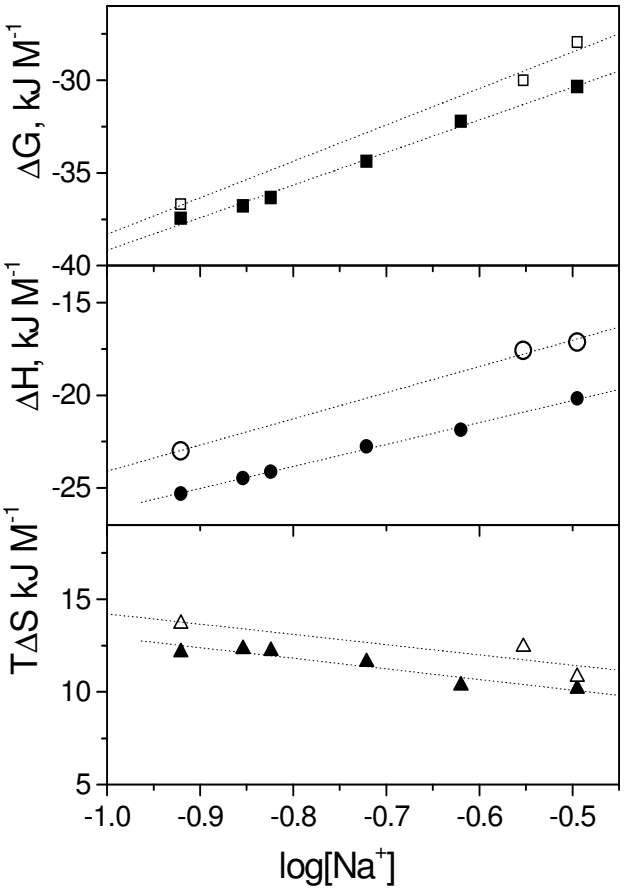


Figure 27. Dependence of observed thermodynamic quantities ΔG (panel A), ΔH (panel B) and TΔS (panel C) on the logarithm of [NaCl] (filled symbols) or [BrCl] (open symbols) added to the buffer. Data were obtained by ITC at 25 °C and pH 6.0. The lines represent linear fits to the data.

7.2.2 Energetic partitioning of the salt effect

Salts belong to the natural environment of proteins and DNA. Not surprisingly, in vitro studies of protein-DNA association have documented modulation of protein-DNA association by salt. In cases where ions are not directly (i.e. specifically) involved in binding and do not site-specifically interact with either of the macromolecules, the increase of ion concentration results in a decrease of the free energy of complex formation. This effect is known as the general salt effect. Although protein DNA-binding domains usually have a large proportion of positively charged side chains, the polyelectrolyte nature of the DNA seems to dominate the observed salt effect. The rigid, negatively charged duplex accumulates a shell of cations. Therefore, disruption of the cationic cloud by the protein molecule dominates the total free energy of the binding process.

Experimental results are commonly treated in terms of the classic counter-cation/limiting law model (CC/LL; [125, 126, 129]). The distribution of ions around the DNA polyanion is modeled by two shells. The inner shell, whose density is relatively independent of the bulk salt concentration, consists of strongly condensed cations. The outer shell represents a Debye–Hückel ion cloud loosely associated with the duplex. Its density depends on the salt

concentration in the bulk. Binding leads to the release of counter-ions that are condensed around the macromolecules according to eqn 45. In a series of classical experiments Record and colleagues investigated the binding of short cationic peptides to DNA [130] and formulated eqs 46 and 47 that hold for many protein–DNA systems [126]. Application of eqn 47 leads to 5.1±0.5 cations expelled from the interface of the Int-DBD-DNA complex. This number is in good agreement with the available structural data. There are seven Arg and Lys side chains contacting DNA in the complex interface, five of which are bonded to phosphate groups and two to nucleotide bases [1, 103]. Phosphate groups are likely to bind Na^+, which are displaced by the positively charged side chains of the protein. At high salt concentration there is less entropy gain from cation displacement, as we have observed.

According to the CC/LL model, weaker binding at higher salt concentration originates entirely from the unfavorable cratic entropy of expelling DNA-associated cations (M^+) into the bulk solution [125, 129]. The enthalpic term of the salt dependent free energy change is believed to be close to zero. This conclusion has been based on the lack of temperature dependence of the binding of poly-lysines and polyamines to DNA measured in different salt concentrations [130]. However, we observe a significant enthalpic contribution to the decrease of the binding free energy, in agreement with results from others [131-134]. The salt-dependent enthalpic change may be attributed to specific binding of anions to the protein, binding of cations to the DNA duplex being uniform and probably unspecific. In the present case, the ΔH dependence is unlikely caused by specific anion binding to the protein, since the salt dependence of ΔH is the same for NaCl and NaBr (slopes in Figure 27). In absolute numbers, ΔH is more negative in the presence of Br^- which could be caused by the more chaothropic nature of the bromide anion. Br^- is expected to interact more strongly with the protein than Cl^-, and, hence, displacement of Br^- will be linked to a larger free energy and a larger enthalpic penalty than displacement of Cl^- (more positive $\Delta\Delta G_{im}$ and less negative $\Delta\Delta G_{ii}$; see eqn 48 and the discussion below). Specific binding of ions could, in principle, shift the pK_a of a proximal ionizable group, thus inducing salt-dependent changes in protonation/deprotonation, which can have a significant enthalpic effect. Still another explanation for enthalpic destabilization by salt is preferential anion-specific and highly temperature dependent accumulation of anions in the vicinity of the protein. We can not completely rule out these possibilities, but deem them unlikely as well. Both effects are temperature dependent and would have been detected by a salt-dependent heat capacity change, yet ΔC_p for Int-DBD binding to DNA is the same in low-salt and high-salt buffers.

Since our results provide no hints in favor of the explanations discussed above, we believe that the observed enthalpic weakening of Int-DBD-DNA binding fits better in the

framework of more recent views on the nature of the salt effect. Direct evaluation of the electrostatic solvation free energy of polyelectrolytes by application of the non-linear Poisson-Boltzmann (PB) equation predicts that the general salt effect opposes binding in terms of both enthalpy and entropy [135-138]. According to the PB model, the counter-ion atmosphere is regarded as a single, continuous layer, whose density decreases gradually in the direction perpendicular to the DNA duplex axis. The salt effect is viewed as a variation in the salt dependent contributions to the free energy, ΔG_S, of charging the DNA [135, 136]:

$$\Delta G_S = \Delta G_{im} + \Delta G_{ii} + \Delta G_{org} \quad (48)$$

The salt-dependent term (ΔG_s) is decomposed into contributions from ion–macromolecule interactions (ΔG_{im}), ion–ion interactions (ΔG_{ii}), and from the cratic entropy associated with ion organization (ΔG_{org}). Both ΔG_{im} and ΔG_{ii} comprise electrostatic (Coulombic) enthalpy and dielectric entropy terms, the latter accounting for the dielectric constant change that is caused by the enhanced water mobility upon disruption of the ion cloud around the DNA. Since the dielectric constant is salt dependent [139] and temperature dependent [137], an enthalpic contribution to ΔG_{im} and ΔG_{ii} follows from the calculations. As to the organizational term, ΔG_{org}, which is small and could be either favorable or unfavorable, it is the sum of the cratic entropy of reorganization of the ion atmosphere and a dielectric entropy term. It has been demonstrated that the favorable dielectric entropy contribution is larger than the cratic entropy of ion release in the total energetic balance of eq 48 [137], [138] .

Altogether, the partial disruption of the ionic cloud around DNA from the binding of the protein produces an unfavorable enthalpic effect as demonstrated by computation [137]. The magnitude of the enthalpic penalty is most certainly modulated by system-specific factors, such as the size and the complementarity of the binding interface, the number of direct protein-DNA ionic bonds, the charge distribution influencing the local electrostatic potential, etc. Exclusion of counterions more distant from the DNA may be entropically less favorable because their motional degrees of freedom are less restricted, while the associated enthalpic effect might still be sizeable because the strong electrostatic field of DNA is long-range. In other words, moving away from the DNA the cratic entropic gain could fall more rapidly with the distance than the Coulombic enthalpic loss. The enthalpy/entropy partitioning of the general salt effect is thus expected to depend on the physical size exclusion away from the DNA duplex, i.e. on the size of the protein domain. The lack of a salt-dependent enthalpic effect in the classical experiments of Record et al. could be due to the much smaller excluded volume and the higher flexibility of penta-lysine in comparison to typical DNA-binding domains. Further experimental and theoretical work is needed to support this interpretation.

7.3 Thermodynamics of complex formation in the presence of glycerol.

7.3.1 Experimental results

The effect of glycerol at [Na^+] = 0.15 M was measured by ITC at 25 °C. Thermodynamic parameters are presented in Table 9 and in Figure 28. Binding affinity is not affected by the lowering of the water activity in 28% glycerol. The insensitivity of the binding free energy to the addition of glycerol is confirmed by kinetic experiments showing negligible variation of $K_A = k_{on}/k_{off}$ in the range from 0 to 29 % glycerol (Table 9 and Figure 28). The lack of free energy changes upon increase of osmotic pressure justifies the use of eq. 47 to calculate the number of cations excluded from the binding interface. However, a constant ΔG of binding hides the compensatory decrease of the binding enthalpy and entropy. Association becomes enthalpically more favorable but is opposed by an unfavorable entropic change in the presence of glycerol. The glycerol-induced changes of ΔH and $T\Delta S$ compensate each other almost perfectly.

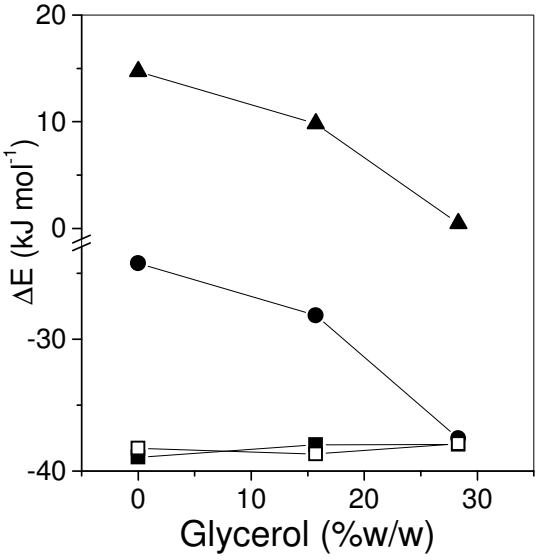


Figure 28. Effect of glycerol on the observed thermodynamic quantities ΔG (open squares-from ITC experiments; solid squares-from stopped flow experiments), ΔH (circles) and $T\Delta S$ (triangles). Data were obtained by ITC at 25 °C and pH 6.0 in the presence of 0.15 M Na^+. The lines represent linear fits to the data.

7.3.2 Binding energetics in the presence of glycerol

Addition of osmolytes has been used to estimate the number of water molecules participating in macromolecular association [140]. Although the thermodynamic foundation of the method has been crizisized, in many cases osmotic stress results fit surprisingly well with structural information [140-142]. We have found evidence that water is intimately involved in Int-DBD-DNA binding [124, 143]. Parsing of the experimental enthalpy and heat capacity changes has led us to the conclusion that the protein-DNA interface is not completely

dehydrated [124]. The NMR structure of the complex identifies cavities large enough to accommodate about 10 water molecules. Molecular dynamics simulations clearly demonstrate the presence of trapped water molecules in the "wet" binding interface, some of them being temporarily disconnected form the bulk solvent [143]. Why then is the binding affinity not influenced by lowering the water activity? The likely explanation is that the Int-DBD complex is rather flexible. On the time scale of our MD simulations (3 ns) water is exchanged rapidly. Most of the water molecules bridging the protein and the DNA form clusters chained to the bulk water. Therefore, the modest lowering of the water activity by 29% glycerol may have been too small to produce a significant osmotic stress. Osmotic pressure considerably affects systems only if a large number of water molecules in stable cavities or in deep crevices are "permanently" disconnected from the bulk solvent.

We observe significant changes in the enthalpy and entropy of association upon addition of glycerol (Figure 28). Enthalpic stabilization and entropic destabilization of the Int-DBD-DNA complex are not easy to explain but could be due in part to hydration effects. Transferring apolar amino acids from water into glycerol-containing solution produces positive changes of enthalpy and entropy. Transfer of polar diglycine leads to negative changes [144]. This implies that burial of apolar surface at the protein-DNA interface in the presence of glycerol favors binding enthalpically (more exothermic) and disfavors binding entropically. Burial of polar surface will counteract in terms of both enthalpy and entropy. Since more apolar surface is shielded from the solvent in the Int-DBD-DNA complex, our results are in line with the general hydration effects.

Transfer enthalpy and entropy contributions (roughly estimated to 1-2 kJ mol^{-1}) [3] can not account for the entire enthalpy/entropy effect observed and there must be other sources for enthalpic stabilization and entropic destabilization induced by glycerol. We propose that addition of glycerol enhances favorable enthalpic interactions between protein and DNA. There is experimental evidence for local and global restriction of the conformational flexibility and thermal motions of proteins in the presence of osmolytes including glycerol as well [145-148]. It is known that glycerol decreases the specific volume and the adiabatic compressibility of native proteins, possibly by causing collapse of voids [149, 150]. Restricted flexibility of the protein-DNA complex in glycerol may intensify the packing interactions, thus increasing the enthalpic component of binding. Along the same line, the observed glycerol-induced "excess" exothermicity of Int-DBD-DNA binding could be interpreted in the light of the more general concept of "enthalpic chelate effect"[4] in associating systems that are stabilized by multi-point weak interactions [151-153]. Recently, D. Williams and colleagues suggested that damping of intermolecular motion due to multi-point binding leads to improved packing with a favorable

enthalpic effect [152]. Clearly, in a cooperative system the opposite is also true: restricted intramolecular thermal motions will increase the enthalpic component of intermolecular contacts. On the other hand, a better packed, more rigid and less fluctuating state is entropically disfavored, the more so in the case of a non-covalently associated complex, for which the vibrational entropy contributions might be considerable [154]. We found that non-covalent leucine zipper dimers are stabilized by enthalpy and destabilized by entropy at low temperatures in the presence of 30% glycerol [155]. Whatever the exact molecular mechanism, it appears that enthalpic stabilization and entropic destabilization around room temperature are the phenomenological thermodynamic signature of the glycerol effect on non-covalent macromolecular systems.

7.4 Kinetics of Int-DBD-DNA binding.

7.4.1 Experimental results

Stopped-flow measurements were performed in the same range of salt concentrations as used in the ITC experiments. The observed half-times of complex formation vary with protein and DNA concentration and the kinetic traces are well described by a simple bimolecular reaction:

$$P + D \underset{k_{off}}{\overset{k_{on}}{\leftrightarrows}} PD$$

Since the half-times of association and dissociation were of similar magnitude under the chosen reaction conditions and less than 100% of complex was formed at equilibrium, the association and dissociation rate constants could be obtained from single reaction traces [156, 157]. The rate of bimolecular association, k_{on}, is substantially reduced by increasing the salt concentration; the decrease is exponential from $(5.2\pm0.5)\times10^{8}$ M^{-1} s^{-1} at 0.12 M [Na^{+}] to $(4.4\pm0.5)\times10^{7}$ M^{-1} s^{-1} at 0.32 M [Na^{+}]. The rates of dissociation, k_{off}, scatter around 15 ± 10 s^{-1}. Plots of $\log k_{on}$ and $\log k_{off}$ against $\log[Na^{+}]$ are linear with slopes of -3.01 ± 0.32 and 0.47 ± 0.38, respectively (Figure 29). The equilibrium constant calculated from kinetic data, $K_{A,kin} = k_{on}/k_{off}$, exhibits the same salt dependence as $K_{A,ITC}$. The slopes in Figure 26 are the same within error. The number of cations displaced from the binding interface is $Z = 5.4\pm0.5$ from the salt dependence of $K_{A,kin}$ and 4.8 ± 0.2 from the salt dependence of K_A^{ITC}. The kinetic analysis reveals that high salt destabilizes the protein-DNA complex by slowing down the rate of association, the

dissociation rate being almost unaffected. We note that K_A^{kin} is consistently higher than K_A^{ITC} (Figure 26). The reason for this discrepancy is not clear and cannot be attributed to temperature dependence of the binding constant [124]. Disparities between K_A obtained by calorimetry or by fluorescence spectroscopy have been reported previously [158, 159].

The rates of association and dissociation were determined also in buffers containing 16 or 29 % glycerol. Both k_{on} and k_{off} decrease with increasing osmolyte concentration (Figure 29 and Table 9). Since they change in parallel, the thermodynamic equilibrium is not affected up to 29% glycerol.

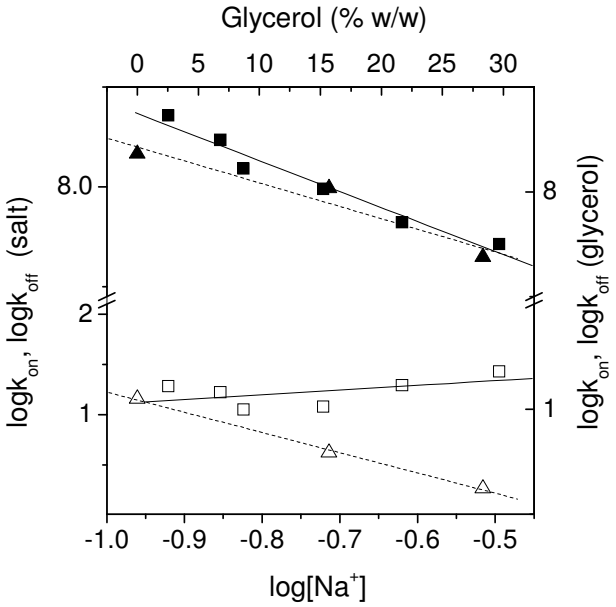


Figure 29. Effect of salt and glycerol on the rates of association and dissociation. Rate constants of association (k_{on}, filled symbols) and dissociation (k_{off}, open symbols) were determined by fluorescence stopped-flow at 7 °C. The lines represent linear fits to the data. Squares and continuous lines (bottom x-axis, left y-axis), data obtained in buffers with varying salt concentration. Triangles and dashed lines (top x-axis, right y-axis), data obtained in glycerol-containing buffer at 0.15 M Na^+.

7.4.2 Solvent effects on the kinetics of association/dissociation

The salt dependence of the rates of association and dissociation of the INT-DBD-DNA complex can be rationalized in terms of electrostatic effects. DNA has a strong negative electrostatic potential and proteins are often positively charged at their DNA-binding face. Since electrostatics is long range, the complementarity of electrostatic fields of protein and DNA speeds up association through diffusion. First, the two molecules are "pulled" together by simple electrostatic attraction. Second, the rotational degrees of freedom are restricted in the initial non-specific encounter complex and the molecules collide in an approximately correct orientation, so that the optimal steric fit required for tight binding is found more rapidly. This effect has been called electrostatic steering [160]. Shielding of charges by salt will slow down the formation of the encounter complex and increase the probability for unproductive encounters. The magnitude

of the steering effect will be modulated in cases where ions are bound or released at the rate-limiting step [161].

Electrostatic interactions are much less important for dissociation because charge-charge (and other polar) interactions suffer a significant energetic penalty from dehydration of the complex interface. This penalty severely diminishes, and quite often overbalances, the favorable Coulombic charge-charge attraction energy. Computational studies predict that non-polar contacts are the major contributor to the stability of protein-DNA complexes and they are much less affected by salt [14]. Therefore, the activation barrier for dissociation is largely dominated by the salt-independent disruption of non-polar interactions. Altogether, the present results confirm that in the absence of specific ion binding effects, salt has a small effect on the rate of dissociation and a large effect on the rate of association.

As to the effect of glycerol on kinetics, the decrease of the association rate is most likely due to the increase of viscosity [162]. The observed decrease of the dissociation rate in glycerol is more difficult to rationalize. In part, it can also be explained by increased viscosity because protein and DNA will diffuse at a lower rate out of the radius of the encounter complex. However, association should be much more affected than dissociation in the low concentration regime of the experiments. We propose a non-negligible contribution from the chelate effect discussed above, which might increase the activation barrier for dissociation of a system stabilized by multi-point weak interactions.

8 THERMODYNAMICS AND KINETICS OF THE INT-DBD MUTANTS BINDING TO DNA

According to the NMR structure of the Int-DBD-DNA complex 13 amino acid side-chains contact the DNA duplex [1]. Alanine scanning mutagenesis has been previously used to quantify the contribution of all side-chain-to-DNA contacts to the binding affinity by using fluorescence quenching and fluorescence anisotropy [103]. Throughout the following discussion we use this data set as a reference to compare the binding energetics determined in our own experiments. Since no structure of a mutant Int-DBD-DNA complex is known, the discussion of structural properties of mutated side chains refers to the structure of the wild-type complex [1]. We performed ITC experiments to measure the binding properties of seven mutants which exhibit the largest effect on ΔG, namely Y40A, K28A, K54A, R55A, R24A, K21A and F38A. The kinetic behavior of four mutants was characterized (Y40A, K21A, R24A, K54A). The changes of all thermodynamic parameters are presented with reference to the wild-type parameters, i.e. $\Delta\Delta J = \Delta J_{mut} - \Delta J_{wt}$, where ΔJ = ΔG, ΔH or TΔS.

8.1 The Y40A mutant

8.1.1 Results

ITC experiments in 50mM phosphate, 100mM NaCl, pH 6.0 showed no observable enthalpic effect at 23°C. Titrations were performed at 4°C and 7°C. Again, there was no measurable heat effect. In theory, the lack of an observable heat effect could be due to mutually compensating heats arising from binding per se and protonation/deprotonation of an ionizable group, which changes it's pK_a upon binding. This scenario was tested by titrations at 5 °C and 23°C in ACES buffer, which has a larger ionization enthalpy than phosphate. We observed heat absorption signals describing a binding isotherm compatible with a binding event of 1:1 stoichiometry (mean 1.006±0.1± SD of 3 experiments). Within experimental error, K_A is temperature-independent between 5 and 23 °C, the mean value being $(1.9 \pm 0.3)\times 10^6$ M^{-1}. Thus, the affinity of Y40A mutant is 3.5 times lower than that of the wild type protein. The apparent enthalpy and entropy changes of association also vary very little with temperature. The heat capacity change, $\Delta C_P = \delta\Delta H_A^{cal}/\delta T$, is negligible (92±2 J K^{-1} mol^{-1} between 5 and 23 °C). Assuming $\Delta H_A^{cal} \sim 0$ kJ mol^{-1} in phosphate buffer and taking the mean value of 13.1±1.0 kJ

mol^{-1} for ΔH_A^{cal} in ACES buffer, we obtain a slope of 0.5 from the plot of ΔH_A^{cal} vs the ionization enthalpy of the used buffers. Hence, net 0.5 protons are released by the buffer and are taken up in the protein-DNA complex.

Kinetic experiments were performed at 5 different concentrations of protein and DNA, keeping their ratio at 1:1. The average of 5 to 10 firings for each concentration of ligand in the range (640 to 1500 nM) was modeled to calculate the rate constants of association and dissociation as described in the Methods section for the wild-type complex. The value of the association rate constant, k_{on}, is $(1.5\pm0.2) \times 10^8$ M^{-1} s^{-1} and is identical within error with the k_{on} for the wild type protein. The dissociation rate constant, $k_{off} = 21.5 \pm 5$ s^{-1}, is higher than for the wild type. Thus K_A calculated from kinetic data is $(7.0\pm0.2) \times 10^6$ M^{-1}.

8.1.2 Discussion

The Y40 side chain is involved in hydrogen bond to cytosine 21 and adenine 20. According to the mutational analysis of Clubb et al. it is the most important contact stabilizing the Int-DBD-DNA complex ($\Delta\Delta G = 7.4\pm0.2$ kJ mol^{-1}). We find a much smaller loss of affinity upon replacement of Tyr 40 by Ala: $\Delta\Delta G = 3.7\pm2$ kJ mol^{-1} from ITC data and $\Delta\Delta G = 1.5\pm1$ kJ mol^{-1} from kinetic data. The binding free energy of the mutant complex is balanced very differently from that of the wild-type Int-DBD-DNA complex. Association is driven almost entirely by favorable entropy changes while the enthalpy is close to zero. Since there is a pronounced energetic effect from a protonation event, it is very difficult to rationalize the observed enthalpic loss and entropic gain in structural terms. Notably, there is a complete loss of ΔC_P combined with a reasonably high affinity. It is possible that water molecules are trapped at the protein-DNA interface in the cavity created when the bulky Tyr side chain is removed. As a consequence, there less non-polar surface is being dehydrated in the complex and ΔCp becomes smaller (since dehydration of non-polar surface has a negative heat capacity effect). However, it is very unlikely that the ~2 kJ K^{-1} mol^{-1} decrease of binding heat capacity change can be attributed solely to hydration effects arising from the removal of a single (even bulky) side chain. One should consider substantial contributions from changes of the overall architecture of the mutant complex. The reason is that the Y40A mutant protein melts at lower temperature and has lower unfolding enthalpy. Therefore, a significant degree of destabilization by removal of the bulky Tyr 40 side chain, and, consequently, some structural distortions at the center of the DNA-binding face are likely. In any case, it is surprising that the dramatic changes of the energetic balance results in only modest loss of binding free energy.

The kinetic analysis reveals that the loss of affinity is caused entirely by the increased rate of dissociation. This is expected because the hydrogen bond and van der Waals packing interactions of the Tyr side chain are short-range and increase the activation energy for dissociation of the complex, thus their removal increases the dissociation rate.

Based on the present results, it is not clear which side chain becomes partly protonated on going from the free to the bound state. The most likely candidate is His 73. This residue is located at the end of the α-helix of Int-DBD, relatively far away from the binding site. However, this α-helix unfolds at its C-term in the wild-type Int-DBD-DNA complex, as demonstrated by NMR. In view of the suggested global changes of the Y40A protein, it is possible that there are differences in the electrostatic environment of His 73 in the mutant complex. Nevertheless, a significant pK_a shift of any of the remaining three Tyr residues can not be completely ruled out.

8.2 The K28A mutant

Lys 28 forms hydrogen bonds to the nucleotide bases G3 and T4 [1] in the protein-DNA interface. Substitution to Ala exhibits the second largest effect on binding affinity for the target duplex: $\Delta\Delta G$ = 6.8 kJ mol^{-1} [103]. Two ITC measurements at 25°C were performed. No heat effect and consequently no binding could be detected. In thermal melting experiments the mutant protein shows a substantial difference between the calorimetrically obtained enthalpy, $\Delta H_{m,cal}$ = 212±10 kJ mol^{-1}, and the van't Hoff enthalpy obtained from CD melting , $\Delta H_{m,vH}$ = 282±10 kJ mol^{-1}. This is usually taken as a sign of an aggregation process taking place. Although thermal melting was in fact reversible after cooling, the quite low calorimetric enthalpy hints at a loss of packing interactions in the folded state. For this reason we abandoned this mutant.

8.3 The K54A mutant

8.3.1 Results

Experiments between 10 and 25 °C revealed relatively high affinity binding with stoichiometry of 0.9±0.1 (mean ± SD of 8 experiments). The binding constant exhibits a small temperature dependence with mean values of $K_A = (1.37\pm0.2)\times10^6$ M^{-1} at 10°C and $K_A = (0.85\pm0.3)\times10^6$ M^{-1} at 25°C. The apparent enthalpy and entropy changes of association vary with temperature. Binding is exothermic at all measured temperatures. The observed enthalpy

changes from –13.2±2.0 kJ mol^{-1} at 10.5°C to –40.7±1.1 kJ mol^{-1} at 25°C (Figure 30A). Almost identical ΔH_A^{cal} was measured in Bis-Tris buffer which rules out protonation contribution (data not shown). The entropy of complex formation is positive below 22°C and negative at higher temperatures (Figure 30A). Linear regression of the $\delta\Delta H_A^{cal}/\delta T$ data yields a $\Delta C_{p,A}$ of –1.9±0.3 kJ K^{-1} mol^{-1}, which is similar to the averaged $\Delta C_{p,A}$ of –2.0±0.2 kJ K^{-1} mol^{-1} for the wild type.

Kinetic experiments were performed at 4 different concentrations in the range 250nM to 1000nM. The kinetic traces were compatible with k_{on} of (42±4) $\times 10^6$ M^{-1} s^{-1}, which is ~3.5 smaller than wild type k_{on} . The rate of dissociation was close to the wild type value of k_{off} = 11.6 ± 2 s^{-1}. The resulting K_A from kinetic data is (3.6±0.8) $\times 10^6$ M^{-1} at 7 °C.

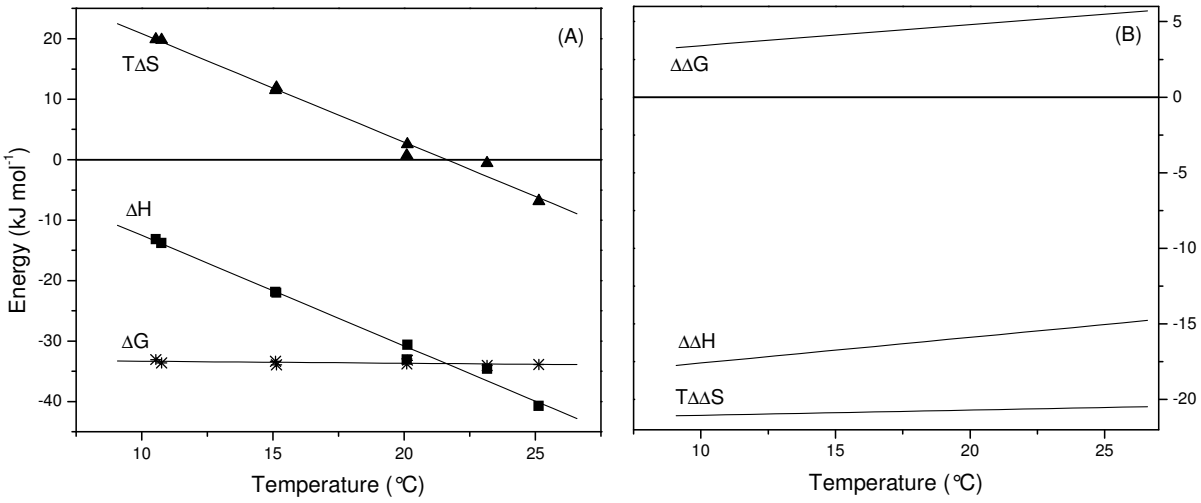


Figure 30. Energetics of DNA association of K54A mutant measured by ITC. (A) Linear fit of $\delta\Delta H_A/\delta T$ yields $\Delta C_{p,A}$ of –1.9±0.3 kJ K^{-1} mol^{-1}. (B) Linear fits of the differential thermodynamic parameters ($\Delta\Delta J = \Delta J_{mut} - \Delta J_{wt}$)

8.3.2 Discussion

The side chain amino-group of K54 is H-bonded to a backbone phosphate of DNA. This H-bond is possible because in the complex DNA is bend and thereby anchored to the C-terminal part of the β-sheet. The importance of the bond is manifested by the six-fold decrease of K_A upon alanine substitution, as measured by fluorescence spectroscopy. We obtain $\Delta\Delta G$ = 5.6 ± 2.7 kJ mol^{-1}, in reasonable agreement with the results of Clubb et al. ($\Delta\Delta G$ = 4.5 kJ mol^{-1}). This modest decrease of affinity results from much larger compensating changes in enthalpy and entropy: $\Delta\Delta H$ = −14.9±3 kJ mol^{-1} and $T\Delta\Delta S$ = −20.5±4kJ mol^{-1} (Figure 30B). From these values, one can conclude that K54 destabilizes the complex enthalpically and stabilizes the

complex entropically. This is surprising because the mutation removes a hydrogen bond (enthalpically favoring the bound state) and is expected to yield free energy gain from elliminating the conformational entropy associated with freezing of the lysine side chain. Since K54 is not much buried in the interface, the thermodynamic profile of the K54A mutant is difficult to be explained by general hydration effects. That hydration is not dominating the observed energetics is supported by the very small heat capacity change of the Lys 54-to-Ala substitution: $\Delta\Delta C_P = 0.1\pm0.4$ kJ mol^{-1} K^{-1}. Possibly, in the absence of the K54 side chain, the DNA duplex is bent to a lesser extent. It is known that distortions of the DNA molecules in protein–DNA complexes (slight base unstacking) is favored entropically and disfavored enthalpically [49].

The kinetic measurements show that the mutation significantly decreases the rate of association and has no effect on the rate of association. Therefore, Lys 54 is important for steering the protein and the DNA toward the encounter complex and for assisting the proper orientation allowing for the optimal fit of the interacting surfaces. This is expected based on general consideration about the role of long-range electrostatic effects. Since the Coloumbic attraction between charges is largely (or completely) outplayed by the unfavorable dehydration of charges, the energy of the K54-DNA H-bond is possibly small and thus removal of this bond does not affect the dissociation step.

8.4 The R55A mutant

8.4.1 Results

The R55A mutant binds DNA with stoichiometry of 0.95±0.1 (mean ± SD of 5 experiments). The substitution of Arg 55 by Ala also weakens binding. The mean K_A is (0.12 ±0.2)×10^6 M^{-1} and is temperature independent within error between 4 and 25 °C. The enthalpy and entropy changes of association vary with temperature and almost completely compensate each other. Binding is endothermic at temperatures below 13°C and exothermic above this temperature. ΔH_A^{cal} changes from –13.9±1.7 kJ mol^{-1} at 4.3°C to –21.2±1.5 kJ mol^{-1} at 25°C (Figure 31A). The entropy of complex formation is positive in the measured temperature interval and is predicted to change sign at around 30°C. Linear regression of the $\delta\Delta H_A^{cal}/\delta T$ data yields a $\Delta C_{p,A}$ of –1.6±0.2 kJ K^{-1} mol^{-1}, as compared to the temperature average of $\Delta C_{p,A}$ = –2.0±0.2 kJ K^{-1} $mol^{-1}$for the wild type complex (Figure 31A). ΔH_A^{cal} equals –14.7±1.3 kJ mol^{-1}

in BisTris buffer at 25°C. The slope of the ΔH_A^{cal} vs ΔH_A^{ion} plot is 0.26. Therefore, it is possible that there is a proton uptake upon complex formation due to the decreased pK_a of a protein side

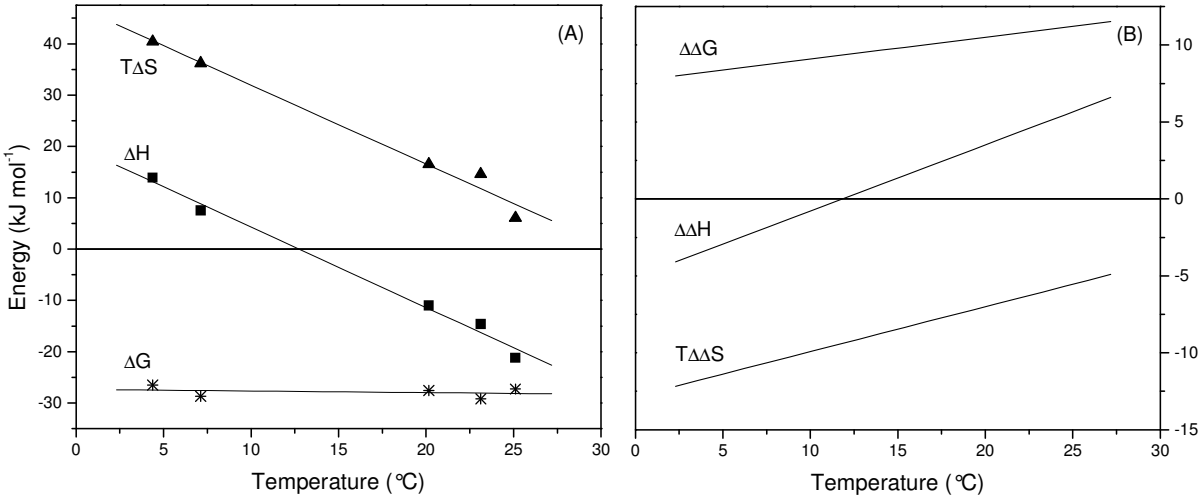


Figure 31. Energetics of DNA association of R55A mutant measured by ITC. (A) Linear fit of $\delta\Delta H_A/\delta T$ yields $\Delta C_{p,A}$ of –1.6±0.2 kJ K^{-1} mol^{-1}. (B) Linear fits of the differential thermodynamic parameters ($\Delta\Delta J = \Delta J_{mut} - \Delta J_{wt}$)

chain, although the precision of the data precludes a definite statement.

8.4.2 Discussion

The position of the Arg 55 side chain in the protein-DNA complex is not well defined and Arg 55 is possibly quite mobile, yet it is important for binding. It contacts DNA in only 10% of the NMR conformers [2]. However, structural analysis reveals that there are no steric restrictions for Arg 55 to form alternating hydrogen bonds with three backbone phosphates. It has been suggested that such a dynamic interaction pattern will be energetically favorable for entropic reasons since the side chain is relatively disordered and fluctuating. Surprisingly, Arg 55 contributes only little to the ΔG of binding according to spectroscopic data, yet the R55A mutation causes the largest decrease of K_A (56 fold) in our experiments. From ITC experiments $\Delta\Delta G$ = 11.3±2.7 kJ mol^{-1}. As shown in Figure 31B the replacement of Arg 55 is penalized by both enthalpy and entropy at around room temperature. While the enthalpic effect of the mutation could be partly attributed to the removed hydrogen bond(s), the origin of entropic destabilization is not obvious. Contributions from conformational changes of the protein can be excluded since the mutation doesn't influence the stability of the protein. Most likely the entropic penalty for substituting the Arg 55 side chain arise from differences in hydration of the

mutant and wild type complexes. The largest decrease in absolute value of the ΔC_P of the binding is observed for the R55A mutant (0.4±0.3 kJ mol^{-1} K^{-1}). Since Arg 55 resides in the extended loop 2 and is located at the periphery of the binding sites it is possible that the interface is opened and "flooded" with water molecules when the Arg 55-DNA contact is not formed. This means that less apolar surface is dehydrated upon binding. Both effects will lead to a lower ΔCp and TΔS.

8.5 The R24A mutant

Arg 24 makes a hydrogen bond with a DNA backbone phosphate group. Abolishing this contact by Ala substitution showed an 8 fold decrease of K_A ($\Delta\Delta G$ = 6.0±0.1 kJ mol^{-1}) in the study of Clubb et al. [103]. Binding was not observable by ITC in the original buffer: no heat effects were measured neither at 7 °C nor at 23° C. A binding event (stoichiometry of 0.84 mol protein per mol of DNA) was clearly detected in buffer containing only 5mM phosphate, pH 6.0. Although the calculated binding constant $(3 \pm 0.6)\times10^6$ M^{-1} is in the range of affinities of most other complexes tested, the reaction was extremely exothermic (ΔH_A^{cal} ~ –130 kJ mol^{-1}) and very unfavorable in entropy (TΔS ~ 90 kJ K^{-1} mol^{-1}). Since there are no reference data with the wild type protein at very low ionic strength, binding of the R24A mutant was not characterized further.

8.6 The K21A mutant

8.6.1 Results

This mutant binds to DNA with K_A = $(0.42\pm0.2)\times10^6$ M^{-1} (average stoichiometry 1.04±0.1), which is temperature independent within error between 4 and 25 °C. Binding is endothermic at temperatures below 10.5°C and becomes exothermic above this point. The ΔH_A^{cal} changes from 10.9±0.5 kJ mol^{-1} at 4.3°C to –23.5±2.2 kJ mol^{-1} at 25°C to (Figure 32A). The apparent enthalpy measured in Bis-Tris buffer at 25°C is –32.1±1.6 kJ mol^{-1}. The dependence of ΔH_A^{cal} on the buffer ionization enthalpy (ΔH_A^{ion}) hints at release of 0.35 protons from the complex into the buffer. The entropy of complex formation is positive in the measured temperature interval and is expected to change sign around 30°C. The heat capacity change is $\Delta C_{p,A}$ = –1.7±0.2 kJ K^{-1} mol^{-1} (Figure32A), 0.3 kJ K^{-1} mol^{-1} lower than $\Delta C_{p,A}$ for wild type.

Kinetic measurements were conducted at 3 different concentrations in the range 750 nM to 3000 nM. The kinetic traces yield $k_{on} = (22 \pm 10) \times 10^6$ M^{-1} s^{-1} and $k_{off} = 75.2 \pm 12$ s^{-1}. Thus

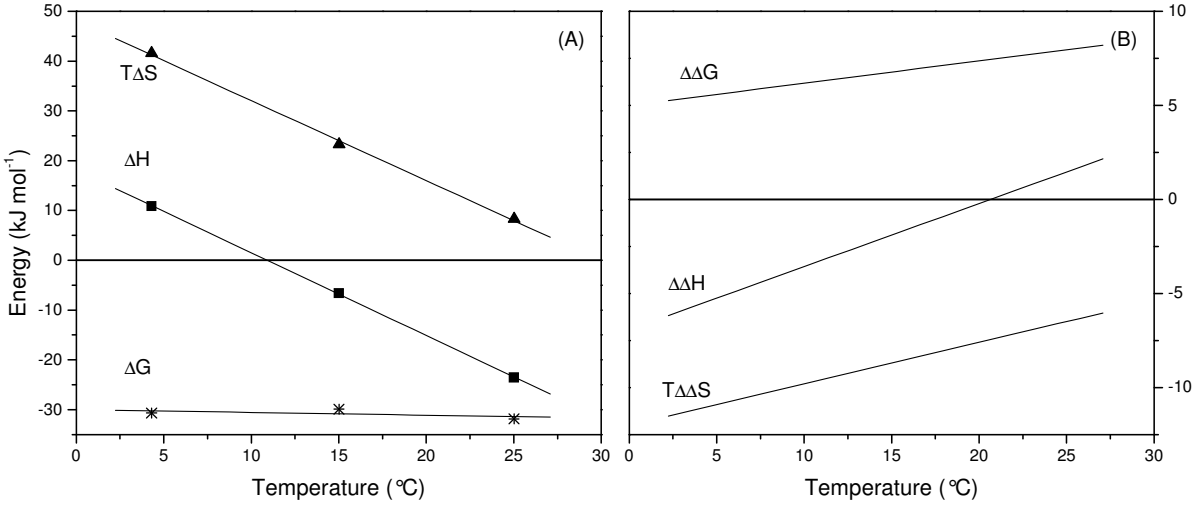


Figure 32. Energetics of K21A mutant -DNA association measured by ITC. (A) Linear fit of $\delta\Delta H_A/\delta T$ yields $\Delta C_{p,A}$ of –1.7±0.2 kJ K^{-1} mol^{-1}. (B) Linear fits of the differential thermodynamic parameters ($\Delta\Delta J = \Delta J_{mut} - \Delta J_{wt}$)

the mutant complex associates at a lower rate and dissocates faster than than the wild type complex. K_A calculated from kinetic data is $(0.24\pm0.1) \times 10^6$ M^{-1}.

8.6.2 Discussion

According to our data the K21A mutation reduces K_A 16 fold (ITC experiments) and 56 fold (kinetic measurements). The loss of free energy is $\Delta\Delta G = 8.0 \pm 2.7$ kJ mol^{-1} (ITC data). The changes in enthalpy and entropy at 25 °C caused by the mutation are $\Delta\Delta H = 1.7\pm2.1$ kJ $mol^{-1}$and $T\Delta\Delta S$ –6.3±3.4 kJ mol^{-1}, respectively (Figure 32B). Depending on the temperature the enthalpic consequences of replacing Lys 21 are either favorable or unfavorable but are relatively small. The loss of binding affinity is dominated by the unfavorable entropic contribution. The change in the solvent accessible surface of Lys 21 in the process of complex formation is –74 A^2, which sets the calculated enthalpic effect of dehydration for the K21A mutation at –53 kJ mol^{-1} while the entropic term amounts to –9.9 kJ mol^{-1}. Considering the mean energy of disruption of a H-bond in vacuum, 40-50 kJ mol^{-1}, the calculated net enthalpic effect of the mutation is on the order of –5 kJ mol^{-1}. Some gain of conformational entropy with an upper limit of 8 kJ mol^{-1} [163] is expected since the contact fixing the Lys 21 side chain is not present in the mutant complex. The net entropic effect of the K21A mutation is therefore between –5 and –2 kJ mol^{-1}

in the rigid body approximation. Thus the predicted contribution of the mutation correlates reasonably with the experimental data. In view of this, the $\Delta\Delta C_P$ of 0.3±0.3 kJ mol^{-1} K^{-1} is difficult to interpret since the buried surface differences between wild-type and mutant complexes are too small to produce a ΔΔCp effect of this magnitude. Lys 21 is located in turn 1, which is pulled toward the duplex in the complex. It is possible that the observed heat capacity decrease in the mutant is due to some relaxation of the strained conformation of turn 1 when the Lys 21-DNA hydrogen bond is removed by mutation.

The decrease of the association rate is expected considering the role of long-range electrostatic effects to formation of the encounter complex. The simultaneous increase of the rate of dissociation is more difficult to explain. A recent computational alanine scanning analysis has indicated a sizeable loss of non-polar energy (van der Waals interaction plus the concomitant changes in non-polar solvation) when Lys 21 is replaced by alanine [15]. Removal of the hydrophobic interactions of the Lys side chain possibly lowers the activation barrier for unbinding.

8.7 The F38A mutant

8.7.1 Results

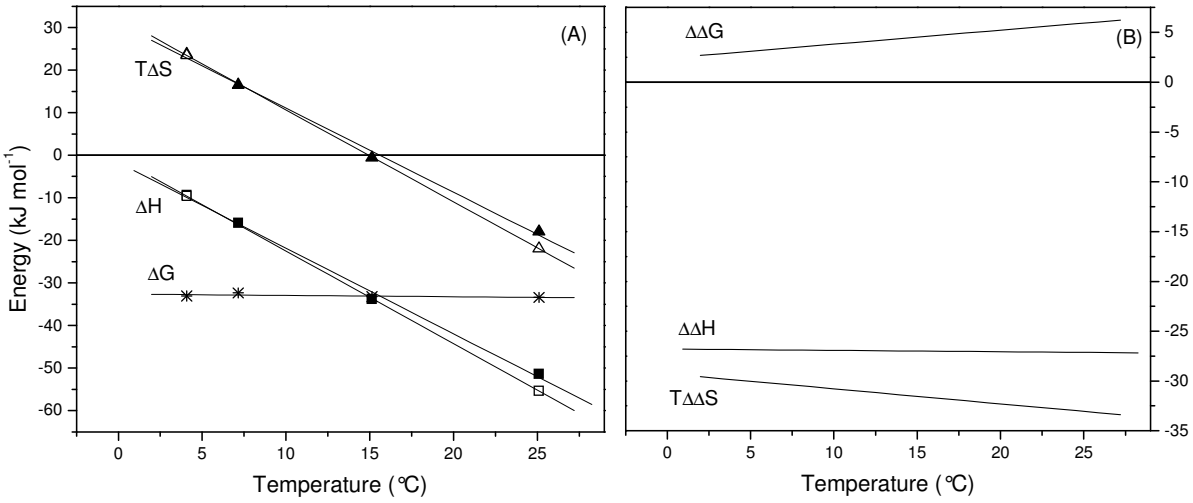


Figure 33. Energetics of DNA association of F38A mutant measured by ITC. (A) Solid symbols represent the ΔH and TΔS without correction for the protonation; open symbols corrected for protonation effect (see the main text); linear fit of $\delta\Delta H_A/\delta T$ (non-corrected) yields $\Delta C_{p,A}$ of –2.0±0.2 kJ K^{-1} mol^{-1}. (B) Linear fits of the non-corrected differential thermodynamic parameters ($\Delta\Delta J = \Delta J_{mut} - \Delta J_{wt}$)

The substitution of Phe38 with Ala did not result in any significant change of K_A in previous experiments [103]. Nevertheless, interaction of Phe 38 with DNA might contribute to binding in terms of enthalpy and entropy. In our ITC experiments the binding constant exhibits a small temperature dependence between $(1.71 \pm 0.3)\times10^6$ M^{-1} at 4.1°C and $(0.71 \pm 0.3)\times10^6$ M^{-1} at 25°C. The binding affinity is thus decreased ($\Delta\Delta G = 6.1 \pm 3$kJ mol^{-1} at 25 °C). Association is exothermic at all measured temperatures. The observed enthalpy ranges from -9.3 ± 0.6 kJ mol^{-1} at 4.1°C to -51.4 ± 3.8 kJ mol^{-1} in the standard phosphate buffer at 25°C (Figure 33A). The enthalpy is -23.7 ± 1.7 kJ mol^{-1} in Bis-Tris buffer at 25°C. It follows that ~1 proton is taken up in the complex upon binding. The entropy of complex formation is positive below 15.8°C and negative at higher temperature. Linear regression of the $\delta\Delta H_A^{cal}/\delta T$ data yields a $\Delta C_{p,A}$ of -2.0 ± 0.2 kJ K^{-1} mol^{-1} (Figure 33A), identical with the temperature average for ΔCp of the wild type complex.

8.7.2 Discussion.

The destabilization of the Int-DBD-DNA complex by removal of the bulky benzene ring of Phe 38 is caused by unfavorable entropic factors, which dominate the stabilizing enthalpy component. The calculated removal of $15A^2$ of aliphatic and $62A^2$ of aromatic surface upon mutation translates into a negligible enthalpic effect (the –11 kJ mol^{-1} gain of dehydration is almost completely counterbalance by the loss of packing interactions) and –8.5 kJ mol^{-1} of entropy ($T\Delta S$). Both estimates can not account for the large $\Delta\Delta H = -30.4$ kJ mol^{-1} and $T\Delta\Delta S = -36.5$ kJ mol^{-1} observed. The lack of heat capacity effect upon substitution of Phe 38 is surprising since less nonpolar and aromatic surface becomes dehydrated in the complex and a decrease of ΔCp is expected. However, removal of the bulky side chain possibly introduces large distortion or a cavity in the protein-DNA interface. If such a cavity is accommodating water, restriction of the thermal fluctuations and vibrational freedom of water molecules would counteract the hydration contribution to ΔC_P. Trapped water might be involved in strong interactions with the protein and/or DNA and this might account also for the large enthalpic and entropic terms of $\Delta\Delta G$ in a way which we do not fully understand.

As in the case of the Y40A mutant, the ionizable group that is responsible for proton uptake is probably Glu 35 based on its proximity to the mutation site. In the case of the F38A mutant the increase of the pK_a of the titrating group is even larger, since in this case binding induces ~100% protonation.

8.8 General discussion

The results presented above clearly indicate that removal of several side chains contacting DNA significantly weaken the binding affinity. However, there is a pronounced quantitative disagreement with the work of Clubb et al. There is no correlation between ΔΔG values determined by the two sets of measurements. For example, the strongest contributor to ΔG in the data set of Clubb et al. is Tyr 40. According to our ITC experiments replacement of Arg 55 is most destabilizing, while Tyr 40 has only a modest destabilizing effect. In clear contradiction to Clubb et al. who demonstrated complete tolerance of the F38A substitution (in, fact this mutation was slightly stabilizing) we find this mutation to be destabilizing by 6 kJ mol^{-1}. The effect of K21A and R55A mutations is more than twice larger from ITC experiments. The reason for these discrepancies is probably methodological. Clubb et al. used a fluorescence anisotropy assay to measure binding of the GST-Int-DBD fusion construct to fluoresceine-labeled DNA duplex extended by five G/C pair on each end. Our experiments were performed with the Int-DBD in isolation and the 13 bp duplex used to solve the NMR structure of the complex. To what extend the different experimental set-ups would have influenced the results is not clear. We note, however, that the affinity measured by fluorescence anisotropy with the chemically modified molecules is ~4 times higher than the affinity of wild type binding (ΔΔG ~3-4 kJ mol^{-1}). This points to some differences in the behavior of the molecules in the two types of assays. On the protein side, a relatively large additional domain (GST, ~25 kDa) is linked to the N-terminus of Int-DBD. Clubb et al. demonstrated that GST alone does not interact with DNA. However, the presence of GST might have importantly influenced the flexibility and dynamics of Int-DBD, the more so in constructs in which the mutations remove not only the freely accessible portion of the side chain but also groups contributing to packing interactions (for example Y40A, F38A, K28A). On the DNA side, the additional G/C pairs could substantially stabilize the flanks of the target sequence. The conformation of the much longer duplex in the protein-bound form might differ from that seen in the complex with the 13 bp duplex. On the other hand, one should also point out that the experimental signals observed by fluorescence and calorimetry have very different physical origins. Discrepant results with regard to K_A have been reported for the two methods. Here, we also observe differences in ΔG values obtained by ITC and from kinetic data from fluorescence stopped-flow experiments.

Analysis of the partitioning of the observed ΔΔG values reveals a very complicated picture. There is no correlation whatsoever between the observed changes in binding affinity and

the magnitude and the balance of the enthalpic and entropic contributions to binding. As a general trend the heat capacity decrement upon binding of the mutants is smaller than for wild type binding. In no case (perhaps with the exception of K21A) the agreement between measured energy terms and the structure-based prediction is reasonably close. Since the predictions use the structure of the wild type protein and take into account only molecular surfaces that are removed by mutation, the discrepancies demonstrate that the rigid body approximation is clearly inadequate to describe binding of Int-DBD to DNA. It appears that some of the mutations cause changes in the stability and conformation of Int-DBD. A clear indication for conformational rearrangements on a larger scale are the observed pKa shift(s) of one (or more) group(s) on going from the free to the bound state. This effect is not seen in the wild type protein and possibly indicates that the mutant protein can "buffer" the destabilizing effect of contacts by conformational adaptation to lower the energetic expenditures for DNA binding.

The kinetic experiments performed so far are not sufficient to make final conclusions (if any are possible) about how mutations influence binding kinetics. The results with Y40A, K54A and K21A agree with the general picture according to which electrostatic interactions can effectively increase the rate of association by steering the molecules into an initial encounter complex: replacement of both Lys 54 and Lys 21 slow down binding, while removal of the uncharged Tyr 40 does not have this effect. On the other hand, stabilization by the base-specific Tyr 40-DNA hydrogen bond comes from slowing down the rate of dissociation. However, as the results with K21A demonstrate, charged side chains can also be important in the dissociation step. The degree to which a given charged side chain will influence the energy of the transition state for binding and unbinding will most certainly depend on local structural details.

9 CONCLUSIONS

The main purpose of the analysis of free protein and DNA was to aid in the interpretation of the association reaction between the Tn916 integrase and its cognate target DNA. For one, it was necessary to know the thermal stability and conformational state of the associating molecules in the range of 5 to 30 °C in which calorimetric and spectroscopic protein-DNA binding experiments were performed. The analysis shows that the protein and the short target DNA duplex are folded up to 30 °C. Second, the heat capacity profile of the protein and the DNA had to be clarified in order to interpret the heat capacity profile of the protein-DNA complex. The DSC experiments demonstrate that the protein accumulates temperature-sensitive structural fluctuations below the onset of the cooperative unfolding transition. The gradual, non-linear structural changes of the protein leading to a minor transition around 30 °C could be noticed only from the heat capacity trace but not from the spectroscopic data. The thermal features of the protein, the steep pre-transitional heat capacity increase of the target DNA, the melting enthalpies of both molecules and their heat capacity changes of unfolding are indispensable to understand the energetics of the protein-DNA association reaction. The present work also reveals temperature dependent conformational dynamics and flexibility in a DNA binding domain. These features may be of general importance to sequence-specific DNA recognition in view of the fact that protein-DNA association reactions are often accompanied by remarkable structural rearrangements.

The present study of the binding of INT-DBD to its cognate DNA is the first thermodynamic analysis of DNA recognition by a three-stranded β-sheet. It is instructive to compare the present thermodynamic parameters with the energetics of association reactions promoted by other more common recognition motifs. Sequence-specific protein binding to DNA has been investigated by calorimetry in several systems and representative data have been discussed [18, 54]. It appears that significant negative heat capacity changes are the only common feature of several protein-DNA association reactions. This is reasonable since protein-DNA complexes bury a large amount of apolar molecular surface. As for other systems, $\Delta C_{p,A}$ of the integrase-DNA complex is large and exceeds by far the expected heat capacity effect of surface dehydration. It seems that values of $\Delta C_{p,A}$ that are too large to originate from hydration effects are a hallmark of specific protein-DNA binding, although there may be exceptions [164]. Interestingly, the Kirchoff plot of ΔH_A versus T is curved. We propose that a nonlinear change

of association enthalpy with temperature is caused by the nonparallel and nonlinear change of the heat capacity C_p of the free components and the complex before the main thermal transition.

Unlike heat capacity changes, the enthalpies and entropies of protein-DNA association are highly variable. For example, DNA binding of the γ cI repressor has a favorable ΔH_A of –100 kJ mol^{-1} and an unfavorable $T\Delta S_A$ of –60 kJ mol^{-1} at 25 °C [44]. Just the opposite is seen for DNA recognition by the TATA-box binding protein, which at 25 °C is opposed by ΔH_A of 125 kJ mol^{-1} and favored by $T\Delta S_A$ of 140 kJ mol^{-1} [45]. As a rule, enthalpic and entropic contributions cancel to a large extent so that the binding free energy is typically in the range of 30–45 kJ mol^{-1} [18]. However, because $\Delta C_{p,A}$ is usually large, the enthalpic and entropic contributions to binding are temperature dependent and my change sign at physiological conditions of temperature. At room temperature, ΔH_A and ΔS_A of the integrase-DNA complex is similar to the enthalpy and entropy change of three other major groove binders [54]: the *Antennapedia* C39S homeodomain, the three zinc finger fragment of TFIIIA, and the methionine repressor dimer MetJ. Only MetJ, a member of the tetrameric ribbon-helix-helix protein family, recognizes the target DNA by β-sheets, similar to integrase Tn916. However, there is little structural homology between MetJ and INT-DBD except that two (MetJ) or three (INT-DBD) β-sheet strands are inserted into the major groove with every second side chain pointing toward the DNA. The energetic similarity of the MetJ-DNA and the integrase-DNA complex could be accidental.

The search for structure-energy relationships in protein-DNA association is greatly hampered by the fact that binding is accompanied by structural rearrangements which are system-specific and difficult to account for. Conformational adaptation is typical for macromolecular recognition. The present study illustrates that adaptation occurs at the cost of introducing conformational disorder in parts of a binding domain. The results emphasize that structure-based methods for the prediction of heat capacity changes (and other thermodynamic parameters) fail when thermal motions and soft vibrational modi of the components are severely altered in the complex. However, we present a useful and promising way to deal with this complication, namely to link ITC measurements of protein-DNA association with the independent DSC analysis of the complex and its free components.

The resolving of the structure of the complex and identification of the protein residues contacting the cognate sequence [1] as well as their significance in terms of ΔG [103] poses the interesting question of deciphering the exact energetic nature of the contacts in enthalpic and entropic terms. For this purpose isothermal titration calorimetry was utilized as the only

available method for simultaneous determination of all the thermodynamic parameters of a binding reaction (ΔG, ΔH, TΔS and n). The results from seven different mutants of INT-DBD suggested that in cases where the stability of the protein was affected this resulted in severe decrease of the binding constant possibly due to general distortion of the overall binding motif architecture (K28 and Y40). In the cases in which the ΔΔG value could be actually measured the abolishing of the respective contacts exhibited various enthalpic and entropic contributions often with magnitudes much larger than the change in the free energy. The energetic partitioning was interpreted in the context of the mutated residues chemical nature (polarity) as well as their position in the binding interface. The contacts participating in the distortion of the DNA showed enthalpy – entropy contributions in line with the known cases of DNA binding proteins which also bend their cognate duplexes. The effect of the mutations on the rate constant was in general negative for both association and dissociation step with pronounced propensity of the polar contacts abolishing to affect mostly the k_{on}.

The influence of salt and addition of the osmolyte glycerol on sequence-specific protein-DNA association was also investigated. The increase of $[Na^+]$ from 0.12 to 0.32 M weakens the binding affinity of Int-DBD for DNA by a factor of 20. Calorimetric experiments show that the decrease in binding free energy is caused by simultaneous unfavorable changes in enthalpy and entropy. We discuss this finding in the light of classical and more recent views about the nature of the general salt effect. Addition of glycerol has no influence on the binding affinity. However, the energetic partitioning of the free energy in terms of enthalpy and entropy is very different in aqueous buffer and in glycerol-containing buffer. We propose that the favorable enthalpic contribution and the unfavorable changes in entropy are a consequence from the general hydration properties of polar and apolar groups, in combination with strengthening of the chelate effect in the presence of glycerol. Addition of salt lowers significantly the rate of protein-DNA association while the rate of dissociation is almost unaffected. Differently, in the presence of glycerol both kinetic constants decrease in parallel. Our results point to complicated thermodynamic mechanisms used by macromolecular complexes to buffer environmental changes.

TABLES

Table 1. Thermodynamic parameters for denaturant-induced unfolding of the 74-residue N-terminal DNA-binding domain of bacteria integrase Tn916 [a]

Temperature	**$[D]_{1/2}$ (M)**	**m_D (kJ mol^{-1} M^{-1})**	**ΔG (kJ mol^{-1})**
Urea			
12.1	3.90 ± 0.04	5.05 ± 0.24	19.8 ± 0.8
16.2	3.70 ± 0.05	4.97 ± 0.24	18.3 ± 0.7
22.0	3.20 ± 0.07	4.78 ± 0.26	15.2 ± 0.5
24.5	2.90 ± 0.04	4.70 ± 0.33	13.7 ± 0.7
27.3	2.60 ± 0.07	4.68 ± 0. 28	12.2 ± 0.4
		mean = 4.84 ± 0.33	
GdmCl			
12.1	2.54 ± 0.02	7.58 ± 0.84	19.2 ± 2.7
16.2	2.36 ± 0.02	7.39 ± 0.67	17.5 ± 2.1
22.0	2.03 ± 0.01	7.51 ± 0.33	15.3 ± 0.8
27.3	1.75 ± 0.02	8.00 ± 0.33	13.9 ± 0.5
		mean = 7.62 ± 0.84	

[a] Experiments performed with 2-5 μM protein in 50 mM Na-phosphate, 100 mM NaCl, pH 6.0, supplemented with increasing concentrations of denaturant. Listed errors were estimated from systematic variation of the slopes of the signal observed in the pre-transitional and post-transitional regions of the traces shown in Figure 6.

Table 2. Observable constants of folding (up to 3.66 mol urea) and unfolding (down to 4.40 mol urea) of the INT-DBD measured with fluorescence stopped-flow with excitation at 280nm and emission cut off 320nm[a].

Folding		**Unfolding**	
Urea M	**lnk_{obs}**	**Urea M**	**lnk_{obs}**
0.63	4.33	4.40	1.07
0.95	4.02	4.93	0.93
1.24	3.73	5.04	0.93
1.57	3.42	5.62	1.02
1.76	3.23	5.99	1.14
1.98	3.03	6.41	1.37
2.23	2.79	6.73	1.56
2.39	2.64	7.31	1.92
2.62	2.43		
2.88	2.19		
3.15	1.95		
3.66	1.53		

[a] maxible possible error of lnK_{obs} is ±5%

Table 3. Thermodynamic parameters for thermal protein unfolding measured by CD, fluorescence and DSC. [a]

	T_m (°C)	ΔH_m (kJ mol^{-1})	ΔC_p (kJ K^{-1} mol^{-1})
CD spectroscopy [b]			
	43.4 (219 nm)	258 ± 25 (219 nm)	
	43.7 (285 nm)	230 ± 25 (285 nm)	
Fluorescence spectroscopy [c]			
	44.2	275 ± 30	
Differential Scanning Calorimetry [d]			
single transition Figure 14B	44.1	238 ± 12 (cal) 248 ± 13 (vH) vH/cal = 1.04 ± 0.07	4.0 ± 0.6 (at 25 °C) 3.7 ± 0.6 (at T_m)
single transition Figure 14C	43.7	255 ± 13 (cal) 236 ± 12 (vH) vH/cal = 0.92 ± 0.07	4.9 ± 0.6 (at 25 °C) 5.8 ± 0.6 (at T_m)
two transitions Figure 14C	34.8 (fit) 44.1 (fit)	22 ± 4 (fit) 243 ± 12 (fit)	1.9 ± 0.1 (fit) 2.6 ± 0.5 (fit)
deconvolution analysis Figure 14D	28 ± 5 (fit) 43.7 (fit)	30 ± 10 (fit) 245 ± 19 (fit)	
Kirchhoff plot [e]			5.7 ± 0.4
mean value	43.8 ± 0.3 [f]	255 ± 18 [f,g]	5.0 ± 0.8

[a] Conditions: 50 mM Na-phosphate, 100 mM NaCl, pH 6.0. [b] Maximal possible error from systematic variation of the slopes in the pre-transitional and post-transitional regions of the melting curves. [c] SD from 3 experiments. [d] Errors are SD of 7 experiments. [e] From Figure 11. [f] Mean value for main transition.

Table 4: Thermodynamic parameters from thermal unfolding of the 74-residue N-terminal DNA-binding domain of bacteria integrase Tn916 in the presence of non-denaturing amounts of urea or GdmCl.

urea (M)	**T_m (°C)**	**ΔH_m (kJ mol^{-1})** [a]
0.36	41.4	236
0.76	39.5	226
0.89	38.5	226
1.50	35.1	204
2.10	29.6	168
2.38	28.0	148
2.82	25.2	141
GdmCl (M)	**T_m (°C)**	**ΔH_m (kJ mol^{-1})** [a]
0.4	41.7	243
0.5	42.0	244
0.8	38.1	230
1.0	36.2	205
1.6	28.2	169

[a] Average error is 10%, as estimated from systematic variation of the slopes in the pre- and post-transition regions of the melting curves (Figure 9).

Table 5. Enthalpy (kJ mol^{-1}) of dissociation and unfolding of the 13 bp DNA duplex measured by DSC and by CD [a]

DSC [b]				CD [c]
ΔH_{grad} [d]	ΔH_{peak} [b]	ΔH_{peak} [e]	ΔH_{vH} [f]	ΔH_{vH}
74 ± 2	332 ± 4	341	335 ± 2	334

[a] Conditions: 50 mM Na-phosphate, 100 mM NaCl, pH 6.0. [b] Mean of 4 experiments performed with 33, 65, 90 and 119 μM duplex DNA. [c] From melting monitored by CD, Figure 19. [d] Enthalpy accumulated as a result of the gradual increase of the heat capacity below the main cooperative transition, as obtained by deconvolution analysis of the excess heat capacity function of Figure 16B. [e] Calculated from the slope of a plot of $1/T_m$ against the duplex DNA concentration (Figure 17). [f] From the analysis of the shape of the heat absorption peak according to eqn 31.

Table 6. Thermodynamic parameters of the association of the integrase DNA binding domain with a 13 bp duplex DNA observed by ITC. [a]

T (°C)	ΔH_A^{cal} (kJ mol^{-1})	$\Delta H_{A,corr}$ [b] (kJ mol^{-1})	ΔG_A [c] (kJ mol^{-1})	$T\Delta S_A$ [d] (kJ mol^{-1})	$T\Delta S_{A,corr}$ [e] (kJ mol^{-1})
4.3	15.7±1.0	15.8	–36.1±2.1	51.8±2.3	51.9
5.6	13.3±1.0	13.6	–35.8±3.4	49.1±3.5	49.4
5.7	10.9±4.4	11.2	–33.2±1.4	44.1±4.6	44.4
7.0	12.2±1.0	12.8	–35.3±3.4	47.5±3.5	48.1
7.2	8.3±0.6	8.9	–38.7±3.6	47.0±3.6	47.7
7.5[f]	11.8±0.6	12.4	–35.4±1.0	47.8±1.5	48.4
8.2	10.0±1.1	10.9	–39.5±1.3	49.5±1.7	50.5
8.5[f]	9.8±0.8	10.7	–37.7±2.0	47.5±1.7	48.4
8.8	8.0±0.8	9.1	–36.8±1.0	44.8±1.3	45.9
9.3	7.6±0.2	8.9	–34.6±1.2	42.2±1.2	43.4
17.6	–10.4±1.0	–5.5	–38.1±3.2	27.7±3.3	32.6
20.1	–15.4±1.0	–8.7	–38.1±3.0	22.7±3.2	29.4
22.1	–19.2±1.5	–11.0	–38.1±3.1	18.9±3.4	27.1
23.1	–30.5±1.9	–21.4	–37.5±3.5	7.0±4.0	16.1
25.1	–25.4±1.3	–14.3	–40.0±4.0	14.6±4.2	25.7
26.0	–34.8±1.7	–22.6	–38.1±2.9	3.3±3.4	15.5
27.1[g]	–40.1±4.2	–26.5	–41.5±4.0	1.4±5.6	15.0
28.1	–40.0±3.5	–24.9	–40.0±4.0	0.0±5.3	14.0
29.1	–38.5±3.3	–21.7	–39.1±4.0	0.6±5.2	17.4
30.0	–47.8±4.0	–26.2	–37.8±1.4	–10.0±4.2	9.7

[a] Experiments performed in standard buffer of pH 6.0. Errors of ΔH_A^{cal} and ΔG_A reflect errors of protein and DNA concentration determination, which were obtained as standard errors of the mean from 3–5 rounds of fitting with either the protein or the DNA concentration as the adjustable parameter of a 1:1 complex. The fitting error was <5% for ΔH_A^{cal} and <10% for K_A. From triplicate experiments, the error of ΔH_A^{cal} was 2–3 kJ mol^{-1} and of ΔG_A 3-4 kJ mol^{-1}. The error $z_{T\Delta S}$ of $T\Delta S_A$ was calculated as $z_{T\Delta S} = \sqrt{(z_{\Delta H})^2 + (z_{\Delta G})^2}$.

[b] $\Delta H_{A,corr}$ is the enthalpy change of a hypothetical rigid body association reaction calculated from eqn 42,43; see the text for details.

[c] From $-RT\ln K_A$.

[d] From $\Delta H_A^{cal} - \Delta G_A$.

[e] From $\Delta H_{A,corr} - \Delta G_A$.

[f] Experiment in HEPES buffer.

[g] Experiment in ACES buffer.

Table 7. Thermodynamic parameters of the association of the mutated integrase DNA binding domain with a 13 bp duplex DNA observed by ITC. [a]

Mutant	T (°C)	ΔH_A^{cal} (kJ mol^{-1})	ΔG_A [b] (kJ mol^{-1})	$T\Delta S_A$ [c] (kJ mol^{-1})	$\Delta C_{P,A}$ (kJ mol^{-1} K^{-1})
K21A	4.3	10.9±0.5	-30.7±0.2	41.6±0.5	-1.7±0.2
	15.0	-6.6±0.7	-29.9±0.2	23.3±0.7	
	25.0	-23.5±2.2	-31.9±0.2	8.3±2.2	
F38A	4.1	-9.3±0.6	-33.1±0.4	23.8±0.7	-2.0±0.2
	7.1	-15.9±0.9	-32.4±0.2	16.5±0.9	
	15.1	-33.8±0.1	-33.2±0.1	-0.6±0.1	
	25.1	-51.4±3.8	-33.4±0.2	-18.0±3.8	
K54A	10.5	-13.2±2.0	-33.1±0.5	19.9±2.1	-1.9±0.3
	10.8	-13.8±2.1	-33.6±0.3	19.8±2.1	
	15.1	-21.9±2.9	-33.4±0.9	11.5±3.0	
	15.1	-22.0±2.9	-33.9±0.3	11.9±2.9	
	20.1	-33.1±2.8	-33.7±0.2	0.7±2.8	
	20.1	-30.6±1.5	-33.2±0.1	2.6±1.5	
	23.2	-34.6±4.2	-34.0±0.3	-0.6±4.2	
	25.1	-40.7±1.1	-33.9±0.2	-6.9±1.1	
R55A	4.4	13.9±1.7	-26.5±0.3	40.4±1.7	-1.6±0.2
	7.1	7.5±0.8	-28.7±0.4	36.2±0.9	
	20.2	-11.0±1.2	-27.5±0.2	16.5±1.2	
	23.1	-14.6±0.4	-29.2±0.1	14.6±0.4	
	25.1	-21.2±1.5	-27.3±0.1	6.1±1.5	

[a] Experiments performed in standard buffer of pH 6.0. Errors of ΔH_A^{cal} and ΔG_A reflect errors of protein and DNA concentration determination, which were obtained as standard errors of the mean from 3-5 rounds of fitting with either the protein or the DNA concentration as the adjustable parameter of a 1:1 complex. The error $z_{T\Delta S}$ of $T\Delta S_A$ was calculated as $z_{T\Delta S} = \sqrt{(z_{\Delta H})^2 + (z_{\Delta G})^2}$.

[b] From $-RT\ln K_A$.

[c] From $\Delta H_A^{cal} - \Delta G_A$

Table 8. Kinetic parameters of the association of the wild type and mutated integrase DNA binding domain with a 13 bp duplex DNA observed by fluorescence stopped-flow. [a]

protein	$k_{on} \times 10^{-6}$ ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	$K_A^{kin}{}_b \times 10^{-6}$ (M^{-1})
Wild type	152±10	11±1	13.5±3
Y40A	150±20	22±5	7±0.2
K21A	22±10	75±12	0.24±0.1
K54A	42±4	12±2	3.6±0.8

[a] Experiments performed in standard buffer of pH 6.0, and temperature 7°C. Errors of k_{on} and k_{off} reflect the maximum deviation of the respective values from the fitting of fluorescence traces recorded at different concentrations

[b] From k_{on} / k_{off}.

Table 9. Thermodynamic and kinetic parameters for the association of Int-DBD with a 13 bp duplex DNA at different salt concentration observed by ITC and fluorescence stopped-flow.[a]

$[Na^+]$	ΔH	TΔS	ΔG	K_A^{ITC} $\times 10^{-6}$	k_{on} $\times 10^{-8}$	k_{off}	K_A^{kin} $\times 10^{-7}$
0.12	−25.3	12.2	−37.5	3.72	5.27	19.2	2.74
0.14	−24.4	12.4	−36.8	2.91	2.95	16.8	1.75
0.15	-24.1	12.2	-36.3	2.35	1.52	11.2	1.36
0.19	−22.8	11.6	−34.4	1.06	0.95	12.0	0.79
0.24	−21.9	10.4	−32.3	0.45	0.44	19.7	0.22
0.32	−20.2	10.2	−30.4	0.21	0.23	26.9	0.99
Glycerol							
15.8	−28.2	9.8	−38.0	1.21	1.04±0.1	6.2±0.8	1.64
28.3	−37.5	0.5	−38.0	1.21	0.49±0.26	4.2±0.27	1.17

[a]All the experiments were conducted ot 25°C, pH 6. Na^+ concentration in the presence of glycerol was 0.15 M. Glycerol concentration in w/w %. ΔH, TΔS and ΔG in units of kJ mol^{-1}; K_A in units of M^{-1}; k_{on} in units of M^{-1} s^{-1}; k_{off} in units of s^{-1}. Maximal possible errors: ΔH, 0.8 kJ mol^{-1}; ΔG, 1.0 kJ mol^{-1}; TΔS, 1.4 kJ mol^{-1}; k_{on} and k_{off}, 10%.

Table 10. Diffrential thermodynamic parameters (X) of the association of the Int-DBD with a 13 bp duplex DNA compared to the wild type protein ($\Delta\Delta J = \Delta J_{mut} - \Delta J_{wt}$)[a] observed by ITC at 25°C.[b]

Mutant	$\Delta\Delta H_A^{cal}$ (kJ mol^{-1})	$\Delta\Delta G_A$ (kJ mol^{-1})	$T\Delta\Delta S_A$ (kJ mol^{-1})	$\Delta\Delta C_{P,A}$ (kJ mol^{-1})
K21A	1.7±2.1	8.0±2.7	-6.3±3.4	0.3±0.3
F38A[c]	-27.2±2.2	6.0±2.7	-33.2±3.5	0±0.3
K54A	-14.9±3	5.6±2.7	-20.5±4	0.1±0.4
R55A	6.0±2.1	11.3±2.7	-5.3±3.4	0.4±0.3

[a] The values for ΔJ_{mut} and ΔJ_{wt} are taken from the linear fit of the respective data at 25°C (except for the $\Delta\Delta C_{P,A}$).
[b] Experiments performed in standard buffer of pH 6.0. Errors of $\Delta\Delta H_A^{cal}$ and $\Delta\Delta G_A$ are calculated as $z_{\Delta\Delta X} = \sqrt{\left(z_{\Delta Xmut}\right)^2 + \left(z_{\Delta Xwt}\right)^2}$, the error of $T\Delta\Delta S$ is calculated as $z_{T\Delta\Delta S} = \sqrt{\left(z_{\Delta\Delta G}\right)^2 + \left(z_{\Delta\Delta H}\right)^2}$.
[c] Values not corrected for the protonation effect

FOOTNOTES

[1] Abbreviations: ACES, N-(2-acetamido)-2-aminoethanesulphonic acid; ASA, solvent accessible surface area; bp, base pair; CD, circular dichroism; DSC, differential scanning calorimetry; ΔG, free energy change; GdmCl, guanidinium hydrochloride; ΔH, enthalpy change; HEPES, N-2-hydroxypiperazine-N'-2-ethanesulphonic acid; ITC, isothermal titration calorimetry; ΔS, entropy change; K_A, equilibrium association constant; k_{on}, association rate constant; k_{off}, dissociation rate constant; k_f, folding rate constant; k_u unfolding rate constant. Subscripts and superscripts to thermodynamic parameters have the following meaning: A, value referring to complex formation from folded components; C, value referring to dissociation of complex into unfolded components; cal, value from direct calorimetric measurement; CD, value from CD melting curve; conf, conformational change; D, DNA; m, value at temperature midpoint of thermal transition; P, protein; rb, rigid body association; vH, value from van't Hoff analysis.

[2] To facilitate reading, the N-terminal fragment 2-74 of integrase Tn916 and the 13 bp duplex target DNA are called "protein" (or INT-DBD: Integrase DNA binding domain) and "DNA", respectively.

[3] The calculation is based on the mean transfer enthalpy and entropy of Ala, Val, Leu, Phe and Trp (to model apolar surface contributions), and diglycine (to model polar surface contributions) normalized per $Å^2$, and the known proportion of polar to apolar surface buried in the Int-DBD-DNA interface. The transfer data from reference (52) were used.

[4] The term "chelate effect" refers to the enhanced coordination properties of polyvalent ligands and has been introduced by coordination chemists. In a more general context, it reflects the phenomenon of positive cooperativity: the sum energetic effect of individual non-covalent bonds is less than the energetic effect driving a multi-point binding process. The classical chelate effect is entropic in nature since the first bond formed eliminates the entropic penalty for formation of a bimolecular complex; subsequent bonds are hence enhanced. D. Williams and colleagues suggested that multi-point binding may cause damping of intermolecular motions. As far as motion of the interacting groups weakens the enthalpy of individual interactions, restriction of motion produces a favorable enthalpic component to binding, i.e. an enthalpic chelate effect.

LIST OF REFERENCES

1. Wojciak, J.M., K.M. Connolly, and R.T. Clubb, *NMR structure of the Tn916 integrase-DNA complex.* Nature Struct. Biol., 1999. **6**(4): p. 366-373.
2. Connolly, K.M., J.M. Wojciak, and R.T. Clubb, *Site-specific DNA binding using a variation of the double stranded RNA binding motif.* Nat Struct Biol, 1998. **5**(7): p. 546-50.
3. Holbrook, J.A., M.W. Capp, R.M. Saecker, and M.T. Record, Jr., *Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices.* Biochemistry, 1999. **38**(26): p. 8409-22.
4. Jelesarov, I. and H.R. Bosshard, *Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition.* J Mol Recognit, 1999. **12**(1): p. 3-18.
5. Jelesarov, I., C. Crane-Robinson, and P.L. Privalov, *The energetics of HMG box interactions with DNA: thermodynamic description of the target DNA duplexes.* J Mol Biol, 1999. **294**(4): p. 981-95.
6. Chalikian, T.V., J. Volker, G.E. Plum, and K.J. Breslauer, *A more unified picture for the thermodynamics of nucleic acid duplex melting: a characterization by calorimetric and volumetric techniques.* Proc Natl Acad Sci U S A, 1999. **96**(14): p. 7853-8.
7. Klump, H., in *Landolt Börnstein Neue Serie Group VII*. 1990, Biophysik Springer: Berlin. p. 241-256.
8. Matthews, B.W., *Protein-DNA interaction. No code for recognition.* Nature, 1988. **335**(6188): p. 294-5.
9. Rhodes, D., J.W. Schwabe, L. Chapman, and L. Fairall, *Towards an understanding of protein-DNA recognition.* Philos Trans R Soc Lond B Biol Sci, 1996. **351**(1339): p. 501-9.
10. Choo, Y. and A. Klug, *Physical basis of a protein-DNA recognition code.* Curr Opin Struct Biol, 1997. **7**(1): p. 117-25.
11. Pabo, C.O. and L. Nekludova, *Geometric analysis and comparison of protein-DNA interfaces: why is there no simple code for recognition?* J Mol Biol, 2000. **301**(3): p. 597-624.
12. Benos, P.V., A.S. Lapedes, and G.D. Stormo, *Is there a code for protein-DNA recognition? Probab(ilistical)ly.* Bioessays, 2002. **24**(5): p. 466-75.
13. Luscombe, N.M. and J.M. Thornton, *Protein-DNA interactions: amino acid conservation and the effects of mutations on binding specificity.* J Mol Biol, 2002. **320**(5): p. 991-1009.
14. Jayaram, B., K. McConnell, S.B. Dixit, A. Das, and D.L. Beveridge, *Free-energy component analysis of 40 protein-DNA complexes: a consensus view on the thermodynamics of binding at the molecular level.* J Comput Chem, 2002. **23**(1): p. 1-14.
15. Gorfe, A.A. and I. Jelesarov, *Energetics of sequence-specific protein-DNA association: computational analysis of integrase Tn916 binding to its target DNA.* Biochemistry, 2003. **42**(40): p. 11568-76.
16. Misra, V.K., J.L. Hecht, A.S. Yang, and B. Honig, *Electrostatic contributions to the binding free energy of the lambdacI repressor to DNA.* Biophys J, 1998. **75**(5): p. 2262-73.
17. Spolar, R.S. and M.T. Record, Jr., *Coupling of local folding to site-specific binding of proteins to DNA.* Science, 1994. **263**(5148): p. 777-84.
18. Jen-Jacobson, L., L.E. Engler, and L.A. Jacobson, *Structural and thermodynamic strategies for site-specific DNA binding proteins.* Structure Fold Des, 2000. **8**(10): p. 1015-23.
19. Dyson, H.J. and P.E. Wright, *Coupling of folding and binding for unstructured proteins.* Curr Opin Struct Biol, 2002. **12**(1): p. 54-60.
20. Dickerson, R.E. and T.K. Chiu, *Helix bending as a factor in protein/DNA recognition.* Biopolymers, 1997. **44**(4): p. 361-403.
21. Schwabe, J.W., *The role of water in protein-DNA interactions.* Curr Opin Struct Biol, 1997. **7**(1): p. 126-34.
22. Nadassy, K., S.J. Wodak, and J. Janin, *Structural features of protein-nucleic acid recognition sites.* Biochemistry, 1999. **38**(7): p. 1999-2017.
23. Nadassy, K., I. Tomas-Oliveira, I. Alberts, J. Janin, and S.J. Wodak, *Standard atomic volumes in double-stranded DNA and packing in protein--DNA interfaces.* Nucleic Acids Res, 2001. **29**(16): p. 3362-76.
24. Luscombe, N.M., S.E. Austin, H.M. Berman, and J.M. Thornton, *An overview of the structures of protein-DNA complexes.* Genome Biol, 2000. **1**(1): p. REVIEWS001.
25. Glover, J.N. and S.C. Harrison, *Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA.* Nature, 1995. **373**(6511): p. 257-61.
26. Schwabe, J.W., L. Chapman, J.T. Finch, and D. Rhodes, *The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements.* Cell, 1993. **75**(3): p. 567-78.

27. Kissinger, C.R., B.S. Liu, E. Martin-Blanco, T.B. Kornberg, and C.O. Pabo, *Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions.* Cell, 1990. **63**(3): p. 579-90.
28. Wolberger, C., A.K. Vershon, B. Liu, A.D. Johnson, and C.O. Pabo, *Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions.* Cell, 1991. **67**(3): p. 517-28.
29. Reddy, C.K., A. Das, and B. Jayaram, *Do water molecules mediate protein-DNA recognition?* J Mol Biol, 2001. **314**(3): p. 619-32.
30. Otwinowski, Z., R.W. Schevitz, R.G. Zhang, C.L. Lawson, A. Joachimiak, R.Q. Marmorstein, B.F. Luisi, and P.B. Sigler, *Crystal structure of trp repressor/operator complex at atomic resolution.* Nature, 1988. **335**(6188): p. 321-9.
31. Schwabe, J.W., L. Chapman, and D. Rhodes, *The oestrogen receptor recognizes an imperfectly palindromic response element through an alternative side-chain conformation.* Structure, 1995. **3**(2): p. 201-13.
32. Allemann, R.K. and M. Egli, *DNA recognition and bending.* Chem Biol, 1997. **4**(9): p. 643-50.
33. Ha, J.H., R.S. Spolar, and M.T. Record, Jr., *Role of the hydrophobic effect in stability of site-specific protein-DNA complexes.* J Mol Biol, 1989. **209**(4): p. 801-16.
34. Lee, B., *Enthalpy-entropy compensation in the thermodynamics of hydrophobicity.* Biophys Chem, 1994. **51**(2-3): p. 271-7; discussion 277-8.
35. Liu, L., C. Yang, and Q.X. Guo, *A study on the enthalpy-entropy compensation in protein unfolding.* Biophys Chem, 2000. **84**(3): p. 239-51.
36. Dunitz, J.D., *Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions.* Chem Biol, 1995. **2**(11): p. 709-12.
37. Sharp, K., *Entropy-enthalpy compensation: fact or artifact?* Protein Sci, 2001. **10**(3): p. 661-7.
38. Cooper, A., C.M. Johnson, J.H. Lakey, and M. Nollmann, *Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions.* Biophys Chem, 2001. **93**(2-3): p. 215-30.
39. Jen-Jacobson, L., *Protein-DNA recognition complexes: conservation of structure and binding energy in the transition state.* Biopolymers, 1997. **44**(2): p. 153-80.
40. Misra, V.K., J.L. Hecht, K.A. Sharp, R.A. Friedman, and B. Honig, *Salt effects on protein-DNA interactions. The lambda cI repressor and EcoRI endonuclease.* J Mol Biol, 1994. **238**(2): p. 264-80.
41. Misra, V.K., K.A. Sharp, R.A. Friedman, and B. Honig, *Salt effects on ligand-DNA binding. Minor groove binding antibiotics.* J Mol Biol, 1994. **238**(2): p. 245-63.
42. Yeh, C.S., F.M. Chen, J.Y. Wang, T.L. Cheng, M.J. Hwang, and W.S. Tzou, *Directional shape complementarity at the protein-DNA interface.* J Mol Recognit, 2003. **16**(4): p. 213-22.
43. Makhatadze, G.I. and P.L. Privalov, *Contribution of hydration to protein folding thermodynamics. I. The enthalpy of hydration.* J Mol Biol, 1993. **232**(2): p. 639-59.
44. Merabet, E. and G.K. Ackers, *Calorimetric analysis of lambda cI repressor binding to DNA operator sites.* Biochemistry, 1995. **34**(27): p. 8554-63.
45. O'Brien, R., B. DeDecker, K.G. Fleming, P.B. Sigler, and J.E. Ladbury, *The effects of salt on the TATA binding protein-DNA interaction from a hyperthermophilic archaeon.* J Mol Biol, 1998. **279**(1): p. 117-25.
46. Kauzmann, W., *Some factors in the interpretation of protein denaturation.* Adv Protein Chem, 1959. **14**: p. 1-63.
47. Finkelstein, A.V. and J. Janin, *The price of lost freedom: entropy of bimolecular complex formation.* Protein Eng, 1989. **3**(1): p. 1-3.
48. Yu, Y.B., P. Lavigne, C.M. Kay, R.S. Hodges, and P.L. Privalov, *Contribution of translational and rotational entropy to the unfolding of a dimeric coiled-coil.* Journal of Physical Chemistry B, 1999. **103**(12): p. 2270-2278.
49. Makhatadze, G.I. and P.L. Privalov, *Energetics of protein structure.* Adv Protein Chem, 1995. **47**: p. 307-425.
50. Baker, B.M. and K.P. Murphy, *Prediction of binding energetics from structure using empirical parameterization.* Methods Enzymol, 1998. **295**: p. 294-315.
51. Luque, I. and E. Freire, *Structure-based prediction of binding affinities and molecular design of peptide ligands.* Methods Enzymol, 1998. **295**: p. 100-27.
52. Cooper, A., *Heat capacity of hydrogen-bonded networks: an alternative view of protein folding thermodynamics.* Biophys Chem, 2000. **85**(1): p. 25-39.
53. Lundback, T., J.F. Chang, K. Phillips, B. Luisi, and J.E. Ladbury, *Characterization of sequence-specific DNA binding by the transcription factor Oct-1.* Biochemistry, 2000. **39**(25): p. 7570-9.

54. Privalov, P.L., I. Jelesarov, C.M. Read, A.I. Dragan, and C. Crane-Robinson, *The energetics of HMG box interactions with DNA: thermodynamics of the DNA binding of the HMG box from mouse sox-5.* J Mol Biol, 1999. **294**(4): p. 997-1013.
55. Berger, C., I. Jelesarov, and H.R. Bosshard, *Coupled folding and site-specific binding of the GCN4-bZIP transcription factor to the AP-1 and ATF/CREB DNA sites studied by microcalorimetry.* Biochemistry, 1996. **35**(47): p. 14984-91.
56. Ladbury, J.E., J.G. Wright, J.M. Sturtevant, and P.B. Sigler, *A thermodynamic study of the trp repressor-operator interaction.* J Mol Biol, 1994. **238**(5): p. 669-81.
57. Tidor, B. and M. Karplus, *The contribution of vibrational entropy to molecular association. The dimerization of insulin.* J Mol Biol, 1994. **238**(3): p. 405-14.
58. Sturtevant, J.M., *Heat capacity and entropy changes in processes involving proteins.* Proc Natl Acad Sci U S A, 1977. **74**(6): p. 2236-40.
59. Morton, C.J. and J.E. Ladbury, *Water-mediated protein-DNA interactions: the relationship of thermodynamics to structural detail.* Protein Sci, 1996. **5**(10): p. 2115-8.
60. Madan, B. and K.A. Sharp, *Hydration heat capacity of nucleic acid constituents determined from the random network model.* Biophys J, 2001. **81**(4): p. 1881-7.
61. Wiseman, T., S. Williston, J.F. Brandts, and L.N. Lin, *Rapid measurement of binding constants and heats of binding using a new titration calorimeter.* Anal Biochem, 1989. **179**(1): p. 131-7.
62. Indyk, L. and H.F. Fisher, *Theoretical aspects of isothermal titration calorimetry.* Methods Enzymol, 1998. **295**: p. 350-64.
63. Sigurskjold, B.W., C.R. Berland, and B. Svensson, *Thermodynamics of inhibitor binding to the catalytic site of glucoamylase from Aspergillus niger determined by displacement titration calorimetry.* Biochemistry, 1994. **33**(33): p. 10191-9.
64. Khalifah, R.G., F. Zhang, J.S. Parr, and E.S. Rowe, *Thermodynamics of binding of the CO2-competitive inhibitor imidazole and related compounds to human carbonic anhydrase I: an isothermal titration calorimetry approach to studying weak binding by displacement with strong inhibitors.* Biochemistry, 1993. **32**(12): p. 3058-66.
65. Freire, E., *Differential scanning calorimetry.* Methods Mol Biol, 1995. **40**: p. 191-218.
66. Privalov, P.L. and S.A. Potekhin, *Scanning microcalorimetry in studying temperature-induced changes in proteins.* Methods Enzymol, 1986. **131**: p. 4-51.
67. Gomez, J., V.J. Hilser, D. Xie, and E. Freire, *The heat capacity of proteins.* Proteins, 1995. **22**(4): p. 404-12.
68. Amend, J.P. and H.C. Helgeson, *Calculation of the standard molal thermodynamic properties of aqueous biomolecules at elevated temperatures and pressures II. Unfolded proteins.* Biophys Chem, 2000. **84**(2): p. 105-36.
69. Privalov, P.L. and G.I. Makhatadze, *Heat capacity of proteins. II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects.* J Mol Biol, 1990. **213**(2): p. 385-91.
70. Xie, D., R. Fox, and E. Freire, *Thermodynamic characterization of an equilibrium folding intermediate of staphylococcal nuclease.* Protein Sci, 1994. **3**(12): p. 2175-84.
71. Viguera, A.R., J.C. Martinez, V.V. Filimonov, P.L. Mateo, and L. Serrano, *Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition.* Biochemistry, 1994. **33**(8): p. 2142-50.
72. Griko, Y.V., G.I. Makhatadze, P.L. Privalov, and R.W. Hartley, *Thermodynamics of barnase unfolding.* Protein Sci, 1994. **3**(4): p. 669-76.
73. Spolar, R.S., J.R. Livingstone, and M.T. Record, Jr., *Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water.* Biochemistry, 1992. **31**(16): p. 3947-55.
74. Darden, T.A., D.M. York, and L.J. Pedersen, *Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems.* J. Chem. phys., 1993. **98**: p. 10098.
75. Brandts, J.F. and L.N. Lin, *Study of Strong to Ultratight Protein Interactions Using Differential Scanning Calorimetry.* Biochemistry, 1990. **29**(29): p. 6927-6940.
76. Litvinovich, S.V. and K.C. Ingham, *Interactions between type III domains in the 110 kDa cell-binding fragment of fibronectin.* J Mol Biol, 1995. **248**(3): p. 611-26.
77. Davis, K.G., S.E. Plyte, S.R. Robertson, A. Cooper, and G.G. Kneale, *Comparison of Pf1 and Fd gene 5 proteins and their single-stranded DNA complexes by NMR spectroscopy and differential scanning calorimetry.* Biochemistry, 1995. **34**(1): p. 148-54.
78. Filimonov, V.V. and V.V. Rogov, *Reversible association of the equilibrium unfolding intermediate of lambda Cro repressor.* J Mol Biol, 1996. **255**(5): p. 767-77.
79. van Holde, K.E., *A hypothesis concerning diffusion-limited protein-ligand interactions.* Biophys Chem, 2002. **101-102**: p. 249-54.

80. Berger, C., S. Weber-Bornhauser, J. Eggenberger, J. Hanes, A. Pluckthun, and H.R. Bosshard, *Antigen recognition by conformational selection.* FEBS Lett, 1999. **450**(1-2): p. 149-53.
81. Durr, E. and H.R. Bosshard, *A monoclonal antibody induces opening of a coiled coil. Global protection of amide protons from H/D exchange decreases by up to 1000-fold in antibody-bound triple-stranded coiled coil.* Eur J Biochem, 1997. **249**(1): p. 325-9.
82. Leder, L., C. Berger, S. Bornhauser, H. Wendt, F. Ackermann, I. Jelesarov, and H.R. Bosshard, *Spectroscopic, calorimetric, and kinetic demonstration of conformational adaptation in peptide-antibody recognition.* Biochemistry, 1995. **34**(50): p. 16509-18.
83. Riggs, A.D., S. Bourgeois, and M. Cohn, *The lac repressor-operator interaction. 3. Kinetic studies.* J Mol Biol, 1970. **53**(3): p. 401-17.
84. Halford, S.E. and M.D. Szczelkun, *How to get from A to B: strategies for analysing protein motion on DNA.* Eur Biophys J, 2002. **31**(4): p. 257-67.
85. Seimiya, M. and Y. Kurosawa, *Kinetics of binding of Antp homeodomain to DNA analyzed by measurements of surface plasmon resonance.* FEBS Lett, 1996. **398**(2-3): p. 279-84.
86. Oda, M., K. Furukawa, A. Sarai, and H. Nakamura, *Kinetic analysis of DNA binding by the c-Myb DNA-binding domain using surface plasmon resonance.* FEBS Lett, 1999. **454**(3): p. 288-92.
87. Hart, D.J., R.E. Speight, M.A. Cooper, J.D. Sutherland, and J.M. Blackburn, *The salt dependence of DNA recognition by NF-kappaB p50: a detailed kinetic analysis of the effects on affinityand specificity.* Nucleic Acids Res, 1999. **27**(4): p. 1063-9.
88. Argos, P., A. Landy, K. Abremski, J.B. Egan, E. Haggard-Ljungquist, R.H. Hoess, M.L. Kahn, B. Kalionis, S.V. Narayana, L.S. Pierson, 3rd, and et al., *The integrase family of site-specific recombinases: regional similarities and global diversity.* Embo J, 1986. **5**(2): p. 433-40.
89. Clewell, D.B., S.E. Flannagan, and D.D. Jaworski, *Unconstrained bacterial promiscuity: the Tn916-Tn1545 family of conjugative transposons.* Trends Microbiol, 1995. **3**(6): p. 229-36.
90. Esposito, D. and J.J. Scocca, *The integrase family of tyrosine recombinases: evolution of a conserved active site domain.* Nucleic Acids Res, 1997. **25**(18): p. 3605-14.
91. Jaworski, D.D., S.E. Flannagan, and D.B. Clewell, *Analyses of traA, int-Tn, and xis-Tn mutations in the conjugative transposon Tn916 in Enterococcus faecalis.* Plasmid, 1996. **36**(3): p. 201-8.
92. Lu, F. and G. Churchward, *Conjugative transposition: Tn916 integrase contains two independent DNA binding domains that recognize different DNA sequences.* Embo J, 1994. **13**(7): p. 1541-8.
93. Poyart-Salmeron, C., P. Trieu-Cuot, C. Carlier, and P. Courvalin, *Molecular characterization of two proteins involved in the excision of the conjugative transposon Tn1545: homologies with other site-specific recombinases.* Embo J, 1989. **8**(8): p. 2425-33.
94. Rudy, C., K.L. Taylor, D. Hinerfeld, J.R. Scott, and G. Churchward, *Excision of a conjugative transposon in vitro by the Int and Xis proteins of Tn916.* Nucleic Acids Res, 1997. **25**(20): p. 4061-6.
95. Salyers, A.A., N.B. Shoemaker, A.M. Stevens, and L.Y. Li, *Conjugative transposons: an unusual and diverse set of integrated gene transfer elements.* Microbiol Rev, 1995. **59**(4): p. 579-90.
96. Scott, J.R. and G.G. Churchward, *Conjugative transposition.* Annu Rev Microbiol, 1995. **49**: p. 367-97.
97. Ramakrishnan, V. and S.W. White, *The structure of ribosomal protein S5 reveals sites of interaction with 16S rRNA.* Nature, 1992. **358**(6389): p. 768-71.
98. Burd, C.G. and G. Dreyfuss, *Conserved structures and diversity of functions of RNA-binding proteins.* Science, 1994. **265**(5172): p. 615-21.
99. Bycroft, M., S. Grunert, A.G. Murzin, M. Proctor, and D. St Johnston, *NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of ribosomal protein S5.* Embo J, 1995. **14**(14): p. 3563-71.
100. Kharrat, A., M.J. Macias, T.J. Gibson, M. Nilges, and A. Pastore, *Structure of the dsRNA binding domain of E. coli RNase III.* Embo J, 1995. **14**(14): p. 3572-84.
101. Flick, K.E., M.S. Jurica, R.J. Monnat, Jr., and B.L. Stoddard, *DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-PpoI.* Nature, 1998. **394**(6688): p. 96-101.
102. Allen, M.D., K. Yamasaki, M. Ohme-Takagi, M. Tateno, and M. Suzuki, *A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA.* Embo J, 1998. **17**(18): p. 5484-96.
103. Connolly, K.M., U. Ilangovan, J.M. Wojciak, M. Iwahara, and R.T. Clubb, *Major groove recognition by three-stranded beta-sheets: affinity determinants and conserved structural features.* J Mol Biol, 2000. **300**(4): p. 841-56.
104. Marky, L.A. and K.J. Breslauer, *Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves.* Biopolymers, 1987. **26**: p. 1601-1620.
105. Milla, M.E. and R.T. Sauer, *P22 Are Repressor - Folding Kinetics of a Single-Domain, Dimeric Protein.* Biochemistry, 1994. **33**(5): p. 1125-1133.

106. Plotnikov, V.V., J.M. Brandts, L.N. Lin, and J.F. Brandts, *A new ultrasensitive scanning calorimeter.* Anal Biochem, 1997. **250**(2): p. 237-44.
107. Freire, E. and R.L. Biltonen, *Statistical mechanical deconvolution of thermal transitions in macromolecules. I. Theory and application to homogeneous systems.* Biopolymers, 1978. **17**: p. 463-479.
108. Freire, E., *Statistical thermodynamic analysis of the heat capacity function associated with protein folding-unfolding transitions.* Comments Mol. Cell. Biophys., 1989. **6**: p. 123-140.
109. Freire, E., *Thermal denaturation methods in the study of protein folding.* Methods Enzymol, 1995. **259**: p. 144-68.
110. Karshikoff, A. and R. Ladenstein, *Proteins from thermophilic and mesophilic organisms essentially do not differ in packing.* Protein Eng, 1998. **11**(10): p. 867-72.
111. Makhatadze, G.I., V.N. Medvedkin, and P.L. Privalov, *Partial molar volumes of polypeptides and their constituent groups in aqueous solution over a broad temperature range.* Biopolymers, 1990. **30**(11-12): p. 1001-10.
112. Hubbard, S.J. and J.M. Thornton, *NACCESS, Computer Program, Department of Biochemistry and Molecular Biology, University College, London.* 1993.
113. Murphy, K.P. and E. Freire, *Thermodynamics of structural stability and cooperative folding behavior in proteins.* Adv.Protein Chem., 1992. **43**: p. 313-361.
114. Wilson, R.W. and J.A. Schellman, *The flow linear dichroism of DNA: comparison with the bead-spring theory.* Biopolymers, 1978. **17**(5): p. 1235-48.
115. Myers, J.K., C.N. Pace, and J.M. Scholtz, *Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding.* Protein Sci, 1995. **4**(10): p. 2138-48.
116. Perczel, A., K. Park, and G.D. Fasman, *Deconvolution of the circular dichroism spectra of proteins: the circular dichroism spectra of the antiparallel beta-sheet in proteins.* Proteins, 1992. **13**(1): p. 57-69.
117. Battiste, J.L., R. Li, and C. Woodward, *A highly destabilizing mutation, G37A, of the bovine pancreatic trypsin inhibitor retains the average native conformation but greatly increases local flexibility.* Biochemistry, 2002. **41**(7): p. 2237-45.
118. Zavodszky, P., J. Kardos, Svingor, and G.A. Petsko, *Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins.* Proc Natl Acad Sci U S A, 1998. **95**(13): p. 7406-11.
119. Ladbury, J.E., J.M. Sturtevant, and N.B. Leontis, *The thermodynamics of formation of a three-strand, DNA three-way junction complex.* Biochemistry, 1994. **33**(22): p. 6828-33.
120. Rouzina, I. and V.A. Bloomfield, *Heat capacity effects on the melting of DNA. 1. General aspects.* Biophys J, 1999. **77**(6): p. 3242-51.
121. Milev, *The energetics of sequence-specific binding of Tn916 integrase DNA-binding domain (INT-DBD) to its cognate target site: 1. Biophysical Characterization and Conformational Stability of INT-DBD and DNA duplex.* submitted, 2002.
122. Carra, J.H. and P.L. Privalov, *Energetics of folding and DNA binding of the MAT alpha 2 homeodomain.* Biochemistry, 1997. **36**(3): p. 526-35.
123. Baldwin, R.L., *Temperature dependence of the hydrophobic interaction in protein folding.* Proc Natl Acad Sci U S A, 1986. **83**(21): p. 8069-72.
124. Milev, S., A.A. Gorfe, A. Karshikoff, R.T. Clubb, H.R. Bosshard, and I. Jelesarov, *Energetics of sequence-specific Protein-DNA association: Binding of integrase Tn916 to its target DNA.* Biochemistry, 2003. **42**(12): p. 3481-3491.
125. Manning, G.S., *The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides.* Q Rev Biophys, 1978. **11**(2): p. 179-246.
126. Record, M.T., Jr., J.H. Ha, and M.A. Fisher, *Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA.* Methods in Enzymology., 1991. **208**: p. 291-343.
127. Dragan, A.I., J.R. Liggins, C. Crane-Robinson, and P.L. Privalov, *The energetics of specific binding of AT-hooks from HMGA1 to target DNA.* Journal of Molecular Biology, 2003. **327**(2): p. 393-411.
128. Olmsted, M.C., J.P. Bond, C.F. Anderson, and M.T. Record, *Grand-Canonical Monte-Carlo Molecular and Thermodynamic Predictions of Ion Effects on Binding of an Oligocation (L(8+)) to the Center of DNA Oligomers.* Biophysical Journal, 1995. **68**(2): p. 634-647.
129. Record, M.T., Jr., M.L. Lohman, and P. De Haseth, *Ion effects on ligand-nucleic acid interactions.* J Mol Biol, 1976. **107**(2): p. 145-58.
130. Lohman, T.M., P.L. Dehaseth, and M.T. Record, *Pentalysine-Deoxyribonucleic Acid Interactions - a Model for the General Effects of Ion Concentrations on the Interactions of Proteins with Nucleic-Acids.* Biochemistry, 1980. **19**(15): p. 3522-3530.
131. Lohman, T.M., L.B. Overman, M.E. Ferrari, and A.G. Kozlov, *A highly salt-dependent enthalpy change for Escherichia coli SSB protein-nucleic acid binding due to ion-protein interactions.* Biochemistry, 1996. **35**(16): p. 5272-9.

132. Lundback, T. and T. Hard, *Salt dependence of the free energy, enthalpy, and entropy of nonsequence specific DNA binding.* Journal of Physical Chemistry, 1996. **100**(44): p. 17690-17695.
133. Oda, M., K. Furukawa, K. Ogata, A. Sarai, and H. Nakamura, *Thermodynamics of specific and non-specific DNA binding by the c-Myb DNA-binding domain.* Journal of Molecular Biology, 1998. **276**(3): p. 571-590.
134. Bergqvist, S., M.A. Williams, R. O'Brien, and J.E. Ladbury, *Heat capacity effects of water molecules and ions at a protein-DNA interface.* Journal of Molecular Biology, 2004. **336**(4): p. 829-842.
135. Misra, V.K., J.L. Hecht, K.A. Sharp, R.A. Friedman, and B. Honig, *Salt Effects on Protein-DNA Interactions - the Lambda-Ci Repressor and Ecori Endonuclease.* Journal of Molecular Biology, 1994. **238**(2): p. 264-280.
136. Sharp, K.A., *Polyelectrolyte electrostatics: Salt dependence, entropic, and enthalpic contributions to free energy in the nonlinear Poisson-Boltzmann model.* Biopolymers, 1995. **36**: p. 227-243.
137. Sharp, K.A., R.A. Friedman, V. Misra, J. Hecht, and B. Honig, *Salt effects on polyelectrolyte-ligand binding: comparison of Poisson-Boltzmann, and limiting law/counterion binding models.* Biopolymers, 1995. **36**(2): p. 245-62.
138. Fogolari, F., A.H. Elcock, G. Esposito, P. Viglino, J.M. Briggs, and J.A. McCammon, *Electrostatic effects in homeodomain-DNA interactions.* Journal of Molecular Biology, 1997. **267**(2): p. 368-381.
139. Hasted, J.B., D.M. Ritson, and C.H. Collie, *Dielectric Properties of Aqueous Ionic Solutions .1.2.* Journal of Chemical Physics, 1948. **16**(1): p. 1-21.
140. Parsegian, V.A., R.P. Rand, and D.C. Rau, *Osmotic stress, crowding, preferential hydration, and binding: A comparison of perspectives.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(8): p. 3987-3992.
141. Timasheff, S.N. and T. Arakawa, *Mechanism of Protein Precipitation and Stabilization by Co-Solvents.* Journal of Crystal Growth, 1988. **90**(1-3): p. 39-46.
142. Timasheff, S.N., *Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(15): p. 9721-9726.
143. Gorfe, A.A., A. Caflisch, and I. Jelesarov, *The role of flexibility and hydration on the sequence-specific DNA recognition by the Tn916 integrase protein: a molecular dynamics analysis.* Journal of Molecular Recognition, 2004. **17**(2): p. 120-131.
144. Gekko, K., *Enthalpy and Entropy of Transfer of Amino-Acids and Diglycine from Water to Aqueous Polyol Solutions.* Journal of Biochemistry, 1981. **90**(6): p. 1643-1652.
145. Bizzarri, A.R. and S. Cannistraro, *Solvent Modulation of the Structural Heterogeneity in Feiii Myoglobin Samples - a Low-Temperature Epr Investigation.* European Biophysics Journal with Biophysics Letters, 1993. **22**(4): p. 259-267.
146. Butler, S.L. and J.J. Falke, *Effects of protein stabilizing agents on thermal backbone motions: A disulfide trapping study.* Biochemistry, 1996. **35**(33): p. 10595-10600.
147. Schlyer, B.D., D.G. Steel, and A. Gafni, *Long time-scale probing of the protein globular core using hydrogen-exchange and room temperature phosphorescence.* Biochemical and Biophysical Research Communications, 1996. **223**(3): p. 670-674.
148. Kendrick, B.S., B.S. Chang, T. Arakawa, B. Peterson, T.W. Randolph, M.C. Manning, and J.F. Carpenter, *Preferential exclusion of sucrose from recombinant interleukin-1 receptor antagonist: Role in restricted conformational mobility and compaction of native state.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(22): p. 11917-11922.
149. Priev, A., A. Almagor, S. Yedgar, and B. Gavish, *Glycerol decreases the volume and compressibility of protein interior.* Biochemistry, 1996. **35**(7): p. 2061-2066.
150. Almagor, A., A. Priev, G. Barshtein, B. Gavish, and S. Yedgar, *Reduction of protein volume and compressibility by macromolecular cosolvents: dependence on the cosolvent molecular weight.* Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology, 1998. **1382**(1): p. 151-156.
151. Hunter, C.A. and S. Tomas, *Cooperativity, partially bound states, and enthalpy-entropy compensation.* Chemistry & Biology, 2003. **10**(11): p. 1023-1032.
152. Williams, D.H., E. Stephens, and M. Zhou, *Ligand binding energy and catalytic efficiency from improved packing within receptors and enzymes.* Journal of Molecular Biology, 2003. **329**(2): p. 389-399.
153. Williams, D.H., E. Stephens, and M. Zhou, *How can enzymes be so efficient?* Chemical Communications, 2003(16): p. 1973-1976.
154. Tidor, B. and M. Karplus, *The Contribution of Vibrational Entropy to Molecular Association - the Dimerization of Insulin.* Journal of Molecular Biology, 1994. **238**(3): p. 405-414.
155. Durr, E. and I. Jelesarov, *Thermodynamic analysis of cavity creating mutations in an engineered leucine zipper and energetics of glycerol-induced coiled coil stabilization.* Biochemistry, 2000. **39**(15): p. 4472-4482.

156. Wendt, H., L. Leder, H. Harma, I. Jelesarov, A. Baici, and H.R. Bosshard, *Very rapid, ionic strength-dependent association and folding of a heterodimeric leucine zipper.* Biochemistry, 1997. **36**(1): p. 204-213.
157. Kozlov, A.G. and T.M. Lohman, *Stopped-flow studies of the kinetics of single-stranded DNA binding and wrapping around the Escherichia coli SSB tetramer.* Biochemistry, 2002. **41**(19): p. 6032-6044.
158. Jelesarov, I. and H.R. Bosshard, *Thermodynamics of Ferredoxin Binding to Ferredoxin-Nadp(+) Reductase and the Role of Water at the Complex Interface.* Biochemistry, 1994. **33**(45): p. 13321-13328.
159. Lundback, T. and J.E. Ladbury, *Inconsistencies in binding affinities measured using different techniques.* Biophys. J., 1998. **74**: p. A40-A40.
160. von Hippel, P.H. and O.G. Berg, *Facilitated target location in biological systems.* J. Biol. Chem., 1989. **264**: p. 675-678.
161. Petri, V., M. Hsieh, E. Jamison, and M. Brenowitz, *DNA sequence-specific recognition by the Saccharomyces cerevisiae "TATA" binding protein: Promoter-dependent differences in the thermodynamics and kinetics of binding.* Biochemistry, 1998. **37**(45): p. 15842-15849.
162. Schreiber, G. and A.R. Fersht, *Rapid, electrostatically assisted association of proteins.* Nature Structural Biology, 1996. **3**(5): p. 427-431.
163. Pickett, S.D. and M.J. Sternberg, *Empirical scale of side-chain conformational entropy in protein folding.* J. Mol. Biol., 1993. **231**(3): p. 825-839.
164. Gonzalez, M., S. Weiler, J.A. Ferretti, and A. Ginsburg, *The vnd/NK-2 Homeodomain: Thermodynamics of Reversible Unfolding and DNA Binding for Wild-Type and with Residue Replacements H52R and H52R/T56W in Helix III.* Biochemistry, 2001. **40**: p. 4923-4931.

ACKNOWLEDGEMENTS

I would like to thank Prof. HR Bosshard for giving me the opportunity to work in his lab and being always very helpful.

I am especially thankful to PD Dr. Ilian Jelesarov - he was always there to answer all my questions and provided guidance at every stage of my work. Also, I would like to thank him and his wife for their friendship and hospitality.

To my former and present colleagues from Prof. Bosshard's group – thank you for your help and friendship.

My special gratitude to Dr. Boris Atanasov – I was lucky to attend his courses at the University of Sofia and to be his Diploma Student.

CURRICULUM VITAE

Family name:	Milev
First name:	Stoyan
Date of birth:	04.02.1976
Nationality:	Bulgarian

Education:

1991 – 1995	Technical College; Burgas - Bulgaria; Diploma in Electro-technics
1995 – 2000	University of Sofia; Sofia – Bulgaria; Masters Degree in Biochemistry
2001 – 2004	University of Zurich – doctoral student, Ph.D studies in Biochemistry

Present position:	Ph.D student in the University of Zurich, Prof. Bosshard group